

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 38.8905 Seconds

(without alignments)  
8225.189 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPGGABPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5879	100.0	1137	10	US-09-902-481A-3
2	5857	99.6	1137	10	US-09-902-481A-4
3	5839	99.3	1153	9	US-09-350-259-3
4	5839	99.3	1153	10	US-09-902-481A-1
5	5839	99.3	1153	10	US-09-891-943-3
6	5839	99.3	1153	14	US-10-144-259-30
7	5839	99.3	1153	14	US-10-207-655-176
8	5836	99.3	1137	10	US-09-902-481A-5
9	5832	99.2	1137	10	US-09-902-481A-6
10	5823.5	99.1	1152	9	US-09-945-265-4
11	3446	58.6	1163	15	US-10-116-275-204
12	3423	58.2	1163	9	US-09-350-259-4
13	3423	58.2	1163	10	US-09-891-943-4
14	3388	57.6	1161	9	US-09-350-259-2
15	3388	57.6	1161	10	US-09-891-943-2

16	3372.5	57.4	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3372.5	57.4	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3213.5	54.7	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3213.5	54.7	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3201	54.4	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3201	54.4	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3192.5	54.3	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3192.5	54.3	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3180	54.1	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3180	54.1	1155	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.4	369	12	US-10-087-192-1212	Sequence 1212, Ap
27	1536.5	26.1	1170	9	US-09-945-265-2	Sequence 2, Appli
28	1532.5	26.1	1170	12	US-10-261-164-1	Sequence 1, Appli
29	1509	25.7	1223	16	US-10-408-765A-295	Sequence 295, App
30	1341.5	22.8	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
33	1128.5	19.2	413	9	US-09-350-259-101	Sequence 101, App
34	1128.5	19.2	413	10	US-09-891-943-101	Sequence 101, App
35	1128	19.2	1179	14	US-10-173-551-2	Sequence 2, Appli
36	1128	19.2	1179	15	US-10-173-551-2	Sequence 2, Appli
37	1087	18.5	1188	15	US-10-291-265-810	Sequence 810, App
38	1085.5	18.5	1151	10	US-09-984-130-103	Sequence 103, App
39	1085.5	18.5	1151	10	US-09-836-353A-103	Sequence 103, App
40	1085.5	18.5	1179	12	US-09-918-715-250	Sequence 250, App
41	1083	18.4	1188	15	US-10-291-265-338	Sequence 338, App
42	1079.5	18.4	1189	10	US-09-984-130-35	Sequence 35, Appl
43	1079.5	18.4	1189	10	US-09-836-353A-35	Sequence 35, Appl
44	1079.5	18.4	1189	12	US-10-262-839-4	Sequence 4, Appli
45	1063	18.1	589	12	US-10-261-164-2	Sequence 2, Appli

#### ALIGNMENTS

#### RESULT 1

US-09-902-481A-3  
; Sequence 3, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RPT/RMS/RNK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; PRIOR FILING DATE: 2001-07-09  
; CURRENT FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-3

Query Match 100.0%; Score 5879; DB 10; Length 1137;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFGQSVVQVQGSRRVVVGAQOEIVAAVNRGSLVQCDYSTGSCBPI	60
Db	1	FNLDTENAMTFQENARGFGQSVVQVQGSRRVVVGAQOEIVAAVNRGSLVQCDYSTGSCBPI	60
Qy	61	RLQVFEAVNNSLGLSLAATTSPPOLLAGCTPVHTCSENTYVKGCLCFGLFGLNLRQOPK	120
Db	61	RLQVFEAVNNSLGLSLAATTSPPOLLAGCTPVHTCSENTYVKGCLCFGLFGLNLRQOPK	120
Qy	121	PFELRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLPFLMQYSEEF	180

Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLKSKTLFSLMOYSSEF 180  
Qy 181 RIHFTKEFQNNPNSLIKPIITQLGRTHATGLRKVVRELFTNTGARKNAFKILFLL 240  
Db 181 RIHFTKEFQNNPNSLIKPIITQLGRTHATGLRKVVRELFTNTGARKNAFKILFLL 240  
Qy 241 TDEKFGDPLGYEDVPELDRGVIRYVLFQGFADFASEKSRQBLNTVASKPRDHVQAN 300  
Db 241 TDEKFGDPLGYEDVPELDRGVIRYVLFQGFADFASEKSRQBLNTVASKPRDHVQAN 300  
Qy 301 NFEALKTQNLREKIPAIETGTTGSSSPFEHMSQEGFSAITSNGLPSTVGSVDWAG 360  
Db 301 NFEALKTQNLREKIPAIETGTTGSSSPFEHMSQEGFSAITSNGLPSTVGSVDWAG 360  
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSVLGAQYQHIGLVAMFR 420  
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSVLGAQYQHIGLVAMFR 420  
Qy 421 QNTGWMESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYVTEQTRGGQSVCP 480  
Db 421 QNTGWMESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYVTEQTRGGQSVCP 480  
Qy 481 PRGORARWQCDVLYGEGQGWGRFGAALTVLGVNGDKLTDVAICAPGEEDNRGAVYLF 540  
Db 481 PRGORARWQCDVLYGEGQGWGRFGAALTVLGVNGDKLTDVAICAPGEEDNRGAVYLF 540  
Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600  
Db 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600  
Qy 601 PVLRVKAIEMFNPREVARNVFCNDQVWKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 601 PVLRVKAIEMFNPREVARNVFCNDQVWKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETKLQLENCEIDPVSPIVLRNLF 720  
Db 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETKLQLENCEIDPVSPIVLRNLF 720  
Qy 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
Db 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
Qy 781 GPREFNVTVVRNDBGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
Db 781 GPREFNVTVVRNDBGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
Qy 841 SGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Db 841 SGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Qy 901 QLELPVKYAVYVMTSHGVSTKYNFTASENTRVMOHQYQVSNLQQRSLPISLVPLVPV 960  
Db 901 QLELPVKYAVYVMTSHGVSTKYNFTASENTRVMOHQYQVSNLQQRSLPISLVPLVPV 960  
Qy 961 RLNQTVIWRBPQVTFSENLSSTCHTERLPSSDPLAELRKAPVNCSTAVQORIQCDIP 1020  
Db 961 RLNQTVIWRBPQVTFSENLSSTCHTERLPSSDPLAELRKAPVNCSTAVQORIQCDIP 1020  
Qy 1021 PFGIQEENFATLKNLSFDWYIKTSHNLLIVSTABILENDSVFTLLPQGAFAVRSQTEF 1080  
Db 1021 PFGIQEENFATLKNLSFDWYIKTSHNLLIVSTABILENDSVFTLLPQGAFAVRSQTEF 1080  
Qy 1081 KVEPPVPNPPLIVGSSVGGLLLLALITAAALYKLGFFRQYKDMSEGGPFGABPQ 1137  
Db 1081 KVEPPVPNPPLIVGSSVGGLLLLALITAAALYKLGFFRQYKDMSEGGPFGABPQ 1137

RESULT 2  
US-09-902-481A-4  
; Sequence 4, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-4  
  
Query Match 99.6%; Score 5857; DB 10; Length 1137;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPQEIIVAAANQRGSLYCCDYSTGSCPEI 60  
Db 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPQEIIVAAANQRGSLYCCDYSTGSCPEI 60  
Qy 61 RLQVPVEAVNNSLGLSLAATTSPOLLACGPTVHTQTCSENTYVKGLCPLFGSNLRQOPK 120  
Db 61 RLQVPVEAVNNSLGLSLAATTSPOLLACGPTVHTQTCSENTYVKGLCPLFGSNLRQOPK 120  
Qy 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLKSKTLFSLMOYSSEF 180  
Db 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKELVSTIMEQLKSKTLFSLMOYSSEF 180  
Qy 181 RIHFTKEFQNNPNSLIKPIITQLGRTHATGLRKVVRELFTNTGARKNAFKILFLL 240  
Db 181 RIHFTKEFQNNPNSLIKPIITQLGRTHATGLRKVVRELFTNTGARKNAFKILFLL 240  
Qy 241 TDGKFGDPLGYEDVPELDRGVIRYVLFQGFADFASEKSRQBLNTVASKPRDHVQAN 300  
Db 241 TDGKFGDPLGYEDVPELDRGVIRYVLFQGFADFASEKSRQBLNTVASKPRDHVQAN 300  
Qy 301 NFEALKTQNLREKIPAIETGTTGSSSPFEHMSQEGFSAITSNGLPSTVGSVDWAG 360  
Db 301 NFEALKTQNLREKIPAIETGTTGSSSPFEHMSQEGFSAITSNGLPSTVGSVDWAG 360  
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSVLGAQYQHIGLVAMFR 420  
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVQSVLGAQYQHIGLVAMFR 420  
Qy 421 QNTGWMESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYVTEQTRGGQSVCP 480  
Db 421 QNTGWMESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYVTEQTRGGQSVCP 480  
Qy 481 PRGORARWQCDVLYGEGQGWGRFGAALTVLGVNGDKLTDVAICAPGEEDNRGAVYLF 540  
Db 481 PRGORARWQCDVLYGEGQGWGRFGAALTVLGVNGDKLTDVAICAPGEEDNRGAVYLF 540  
Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600  
Db 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600  
Qy 601 PVLRVKAIEMFNPREVARNVFCNDQVWKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 601 PVLRVKAIEMFNPREVARNVFCNDQVWKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETKLQLENCEIDPVSPIVLRNLF 720  
Db 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETKLQLENCEIDPVSPIVLRNLF 720  
Qy 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780

Db 721 SLVGTPLSAFGLNRPVLAEDQRLFTALPFFPKKNCNDNICQDDLSITFSFMSDCLVVG 780  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840  
Db 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840  
Qy 841 SGALKSTSCSINHPIPPENSEVTFTNITFDVDSKASLGKLLKXNVTSNNPRNTKTEF 900  
Db 841 SGALKSTSCSINHPIPPENSEVTFTNITFDVDSKASLGKLLKXNVTSNNPRNTKTEF 900  
Qy 901 QLELPVKYAVVMVTSYKLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 960  
Db 901 QLELPVKYAVVMVTSYKLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 960  
Qy 961 RLNQTVIWDREPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020  
Db 961 RLNQTVIWDREPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020  
Qy 1021 FFGIOBEFNATLKGNSLSPDNYIKTSHNHLIIYSTAEIILFNDSDVFTLLPQOGAFVRSQET 1080  
Db 1021 FFGIOBEFNATLKGNSLSPDNYIKTSHNHLIIYSTAEIILFNDSDVFTLLPQOGAFVRSQET 1080  
Qy 1081 KVEPPEVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1137  
Db 1081 KVEPPEVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1137

RESULT 3

US-09-350-259-3  
; Sequence 3, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20020062008A1 Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 09/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 9; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ENLDTENATVFOENARGFGQSVVLOGSVVYVGAQOEIVAAQNSGLVOCYSTGSCRP 60  
Db 17 FNLDTENATVFOENARGFGQSVVLOGSVVYVGAQOEIVAAQNSGLVOCYSTGSCRP 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHOTCSENTYVKGCLFLGSLNRQOQPK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHOTCSENTYVKGCLFLGSLNRQOQPK 136  
Qy 121 FREALGCGQSDIAFLIDGSGIIPHDFRMKELVSTIMEOLKSKTLPFLMOYSEF 180  
Db 137 FREALGCGQSDIAFLIDGSGIIPHDFRMKELVSTIMEOLKSKTLPFLMOYSEF 196

Qy 181 RIHETPKFQNNPNRSLIKPIITOLIGRTHATGIRKVVARELFINITNGARCAKAKILL 240  
Db 197 RIHETPKFQNNPNRSLIKPIITOLIGRTHATGIRKVVARELFINITNGARCAKAKILL 256  
Qy 241 TDGEKFGDPLGYDVIDELEGVIRVVLGPDGAFRSEKSEQLNTWASKPRDHVFOAN 300  
Db 257 TDGEKFGDPLGYDVIDELEGVIRVVLGPDGAFRSEKSEQLNTWASKPRDHVFOAN 316  
Qy 301 NFEALKTQNLQREKIFAIETGOTGSSSSPEHESQSGSAITNSNGELLSTVSGSYDWAG 360  
Db 317 NFEALKTQNLQREKIFAIETGOTGSSSSPEHESQSGSAITNSNGELLSTVSGSYDWAG 376  
Qy 361 GVFLYTSKEKSTFTNMTVRDSDMNDAYLGVAIAAIIILNRVQSLVIGAPRYOHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFTNMTVRDSDMNDAYLGVAIAAIIILNRVQSLVIGAPRYOHIGLVAMFR 436  
Qy 421 QNTGMSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQSVCP 480  
Db 437 QNTGMSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQSVCP 496  
Qy 481 PRGORARWQCDVLYGQGGQPGWGFAGALTIVLGVNDGKLTDAVAGPCEEDNKGAVYLF 540  
Db 497 PRGORARWQCDVLYGQGGQPGWGFAGALTIVLGVNDGKLTDAVAGPCEEDNKGAVYLF 556  
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDQLTMDGLVDLTVGAQSHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDQLTMDGLVDLTVGAQSHVLLRSQ 616  
Qy 601 PVLRAVKALMBENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRSGQTSVVT 660  
Db 617 PVLRAVKALMBENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRSGQTSVVT 676  
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVPVLRNF 720  
Db 677 YDLALDSGRPHSRVAFNETKNSRTRQVGLTQTCETLKLQLPNCIEDPVPVLRNF 736  
Qy 721 SLVGTPLSAFGLNRPVLAEDQRLFTALPFFPKKNCNDNICQDDLSITFSFMSDCLVVG 780  
Db 737 SLVGTPLSAFGLNRPVLAEDQRLFTALPFFPKKNCNDNICQDDLSITFSFMSDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIPPENSEVTFTNITFDVDSKASLGKLLKXNVTSNNPRNTKTEF 900  
Db 857 SGALKSTSCSINHPIPPENSEVTFTNITFDVDSKASLGKLLKXNVTSNNPRNTKTEF 916  
Qy 901 QLELPVKYAVVMVTSYKLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 960  
Db 917 QLELPVKYAVVMVTSYKLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 976  
Qy 961 RLNQTVIWDREPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020  
Db 977 RLNQTVIWDREPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1036  
Qy 1021 FFGIOBEFNATLKGNSLSPDNYIKTSHNHLIIYSTAEIILFNDSDVFTLLPQOGAFVRSQET 1080  
Db 1037 FFGIOBEFNATLKGNSLSPDNYIKTSHNHLIIYSTAEIILFNDSDVFTLLPQOGAFVRSQET 1096  
Qy 1081 KVEPPEVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1137  
Db 1097 KVEPPEVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1153

RESULT 4

US-09-902-481A-1  
; Sequence 1, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimooka, Motomu  
; APPLICANT: Shifman, Julia

APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/FT/MS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (17)...()  
; OTHER INFORMATION:  
US-09-902-481A-1

Query Match 99.3%; Score 5839; DB 10; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
QY 1 FNLDTENAMTFOENARGFQGVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFOENARGFQGVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLPFGSLNLRQPOK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLPFGSLNLRQPOK 136  
QY 121 PFEALRGCPQSDSDIAFLIDGSGSIIPHFRMKELVSTIMEOLKSKTLPFSLMOYSEEF 180  
DB 137 PFEALRGCPQSDSDIAFLIDGSGSIIPHFRMKELVSTIMEOLKSKTLPFSLMOYSEEF 196  
QY 181 RIHFTKEFQNNPNRSLKPIQLGRTHATGLRKVVRELFNITNGARKNAKILFLL 240  
DB 197 RIHFTKEFQNNPNRSLKPIQLGRTHATGLRKVVRELFNITNGARKNAKILFLL 256  
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLFQGDAPFSEKRSQELNLTASKPPRDHVFQAN 300  
DB 257 TDGEKFGDPLGYEDVPELDREGVIRYVLFQGDAPFSEKRSQELNLTASKPPRDHVFQAN 316  
QY 301 NFEALKTQVONLRBEKIFATECTGTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDNAG 360  
DB 317 NFEALKTQVONLRBEKIFATECTGTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDNAG 376  
QY 361 GVFLYTSKEKSTPIINTRVDSQNDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTPIINTRVDSQNDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 436  
QY 421 QNTGMHESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSVCP 480  
DB 437 QNTGMHESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETRGQGVSVCP 496  
QY 481 PRQQRARWQDAVLYEQCPWRFGAALTLDVNGDKLTDVAIGAPCEENRGAIVLP 540  
DB 497 PRQQRARWQDAVLYEQCPWRFGAALTLDVNGDKLTDVAIGAPCEENRGAIVLP 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIYFGQSLGGQDLTMDGLVLTGCAQGHVLLRSQ 600  
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIYFGQSLGGQDLTMDGLVLTGCAQGHVLLRSQ 616  
QY 601 PVLRVKAIEMFNPREVARNVFCNDQVVKRAGEVRVCLHVOKSTRDLRREGQIQSVVT 660  
DB 617 PVLRVKAIEMFNPREVARNVFCNDQVVKRAGEVRVCLHVOKSTRDLRREGQIQSVVT 676  
QY 661 YDLALDSGRPHSAVFNETKNSTRQTVLGLTQTCETLKLQPNCTEDPVSPIVLRNLP 720  
DB 677 YDLALDSGRPHSAVFNETKNSTRQTVLGLTQTCETLKLQPNCTEDPVSPIVLRNLP 736  
QY 721 SLVGTPLSAFGNLRPVLAEDAQLFTALPFPEKNCNDNICODDLSITTFSPMSLDCLVVG 780

DB 737 SLVGTPLSAFGNLRPVLAEDAQLFTALPFPEKNCNDNICODDLSITTFSPMSLDCLVVG 796  
QY 781 GPRFNTVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
DB 797 GPRFNTVTVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
DB 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVVMVTSBGVSTKYLNFTASENTSRVMOHQYOVSNLQGSRLPISLVFLVPV 960  
DB 917 QLELPVKYAVVMVTSBGVSTKYLNFTASENTSRVMOHQYOVSNLQGSRLPISLVFLVPV 976  
QY 961 RLNQTVIWRDPQVTFSENLSSSTCHTKERLPSHSDPLAELRKAPVVCNCSIAVCORIQCDIP 1020  
DB 977 RLNQTVIWRDPQVTFSENLSSSTCHTKERLPSHSDPLAELRKAPVVCNCSIAVCORIQCDIP 1036  
QY 1021 PFGIOEEFNATLKGKLSFDWYIKTSHNHLIVSTABILFNDVSFTLLPQOGAFVRSQTEF 1080  
DB 1037 PFGIOEEFNATLKGKLSFDWYIKTSHNHLIVSTABILFNDVSFTLLPQOGAFVRSQTEF 1096  
QY 1081 KVEPPEVNPPLIVGSSVGGLLILALITAAALYKLGPEKROYKDMMSGGPPGAEPO 1137  
DB 1097 KVEPPEVNPPLIVGSSVGGLLILALITAAALYKLGPEKROYKDMMSGGPPGAEPO 1153  
RESULT 5  
US-09-891-943-3  
; Sequence 3, Application US/09891943  
; Publication No. US20030077278A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20030077278A1 Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/891,943  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-891-943-3

Query Match 99.3%; Score 5839; DB 10; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
QY 1 FNLDTENAMTFOENARGFQGVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFOENARGFQGVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLPFGSLNLRQPOK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLPFGSLNLRQPOK 136  
QY 121 PFEALRGCPQSDSDIAFLIDGSGSIIPHFRMKELVSTIMEOLKSKTLPFSLMOYSEEF 180  
DB 137 PFEALRGCPQSDSDIAFLIDGSGSIIPHFRMKELVSTIMEOLKSKTLPFSLMOYSEEF 196  
QY 181 RIHFTKEFQNNPNRSLKPIQLGRTHATGLRKVVRELFNITNGARKNAKILFLL 240  
DB 197 RIHFTKEFQNNPNRSLKPIQLGRTHATGLRKVVRELFNITNGARKNAKILFLL 256

Qy	241	TDGEKFGDPLGYEDVIELDREGVIRYVLFGDAPRSEKSRQELNVTASKPPDRHVFQAN	300
Db	257	TDGEKFGDPLGYEDVIELDREGVIRYVLFGDAPRSEKSRQELNVTASKPPDRHVFQVN	316
Qy	301	NFEALKTVQNLREKIFAIEGTOTGSGSSPEHEMSQSGFSAATNSNGPLLTSGVSYDWAG	360
Db	317	NFEALKTVQNLREKIFAIEGTOTGSGSSPEHEMSQSGFSAATNSNGPLLTSGVSYDWAG	376
Qy	361	GVFLYTSKEKSTFNIMTRVDSMDNDAYLGAAAILLRNVOSVLGAPRYQHIOLGVAMFR	420
Db	377	GVFLYTSKEKSTFNIMTRVDSMDNDAYLGAAAILLRNVOSVLGAPRYQHIOLGVAMFR	436
Qy	421	QNTGMWESNANVKGTOTGAYFGASLCSVDVDSNSTDLVLIGAPHYEQTEGGGVSVCP	480
Db	437	QNTGMWESNANVKGTOTGAYFGASLCSVDVDSNSTDLVLIGAPHYEQTEGGGVSVCP	496
Qy	481	PRGORARWQCDAVLYGSGQGPWGEFGAALTVLGVONGDKLTDVAIGAPGEDNRGAYL	540
Db	497	PRGORARWQCDAVLYGSGQGPWGEFGAALTVLGVONGDKLTDVAIGAPGEDNRGAYL	556
Qy	541	HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTIVGAQHVLILLRSQ	600
Db	557	HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTIVGAQHVLILLRSQ	616
Qy	601	PVLRVKXIMEFNPREVARNPECNDOVYVKGKEAGEVRVCLHVQKSTRDRLRGEIQSV	660
Db	617	PVLRVKXIMEFNPREVARNPECNDOVYVKGKEAGEVRVCLHVQKSTRDRLRGEIQSV	676
Qy	661	YDLALDSGRPHSRVAVNETKNSTRRQTVLGLTQCTETLKLQLENCIEDPVSPIVLRNF	720
Db	677	YDLALDSGRPHSRVAVNETKNSTRRQTVLGLTQCTETLKLQLENCIEDPVSPIVLRNF	736
Qy	721	SLVGTPLSAPGNLPRVLAEDAQRLEFTALFPPEKPCGNNDNICQDDLSITPSPMSJDL	780
Db	737	SLVGTPLSAPGNLPRVLAEDAQRLEFTALFPPEKPCGNNDNICQDDLSITPSPMSJDL	796
Qy	781	GPREFNTVTVVRNDEGSYRTQVTFPPPLDLSYERKVSITLQNRQSRQSWLLACBSAS	840
Db	797	GPREFNTVTVVRNDEGSYRTQVTFPPPLDLSYERKVSITLQNRQSRQSWLLACBSAS	856
Qy	841	SGALKSTSCSINHIFPENSEVTENITPDDVDSKASLGKLLKANVTSENNMPRTNKT	900
Db	857	SGALKSTSCSINHIFPENSEVTENITPDDVDSKASLGKLLKANVTSENNMPRTNKT	916
Qy	901	QLELPVKYAVVMVTVSHQVSTKYLNFTASENTSRVMOHOYOVSNLQORSILPISLVFL	960
Db	917	QLELPVKYAVVMVTVSHQVSTKYLNFTASENTSRVMOHOYOVSNLQORSILPISLVFL	976
Qy	961	RLNQTIVWDRPQVTFSENLSSTCHTKERLPGSHSDFLAEIRKAPVYVNCIAVCQR	1020
Db	977	RLNQTIVWDRPQVTFSENLSSTCHTKERLPGSHSDFLAEIRKAPVYVNCIAVCQR	1036
Qy	1021	FFGJQBEFNATLKGNLSPDWIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSOT	1080
Db	1037	FFGJQBEFNATLKGNLSPDWIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSOT	1096
Qy	1081	KVEPPEVPNPULPIVGSVGGALLLALITAAALYKLGFFKROYKXDMGEGGPPGAE	1137
Db	1097	KVEPPEVPNPULPIVGSVGGALLLALITAAALYKLGFFKROYKXDMGEGGPPGAE	1153

## RESULT 6

RESOLUTION 6  
US-10-144-259-30

US-10-144-259-30  
: Semence 30. Annotation US/10144259

: Sequence 30, Application US/1011  
: Publication No. US20070109697A7

; Publication No. US200

; GENERAL INFORMATION:

APPLICANT: Arnaout, M. Amin

APPLICANT: Li, Rui

APPLICANT: Xiong, Jian-Ping

; TITLE OF INVENTION: VARIANT

: TITLE OF INVENTION: VARIANTS INTEGRIN POLY-  
: FILE REFERENCE: 00786-548001

QY 841 SGALKSTSCSINHPIPEENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900  
DB 857 SGALKSTSCSINHPIPEENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 916  
QY 901 QLELPVKYAVYMWVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 960  
DB 917 QLELPVKYAVYMWVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 976  
QY 961 RLNOTVIWDRPQVTFSENLSSCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDIP 1020  
DB 977 RLNOTVIWDRPQVTFSENLSSCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDIP 1036  
QY 1021 FPGIOEFPNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEP 1080  
DB 1037 FPGIOEFPNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEP 1096  
QY 1081 KVEPPEVNPPLPLIVGSSVGGLLALLIITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137  
DB 1097 KVEPPEVNPPLPLIVGSSVGGLLALLIITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153  
RESULT 7  
US-10-207-655-176  
; Sequence 176, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069.401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 176  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-207-655-176  
Query Match 99.3%; Score 5839; DB 14; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
QY 1 FNLDTENATFQENARSGQSVVQLGSRVWVGAPQEIIVAAANORGLYQCDYSTGSCBPI 60  
DB 17 FNLDTENATFQENARSGQSVVQLGSRVWVGAPQEIIVAAANORGLYQCDYSTGSCBPI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVKGFLGFLGNSLRQPOK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVKGFLGFLGNSLRQPOK 136  
QY 121 FPEALGCGQEDSDIAFLIDGSGIIPHPFRMKELVSTIMEQLKSKTFLSLMOYSEBF 180  
DB 137 FPEALGCGQEDSDIAFLIDGSGIIPHPFRMKELVSTIMEQLKSKTFLSLMOYSEBF 196  
QY 181 RLHFTKEQNNPNRSLKPIPTOLLGRTHATGLRKVVRELFTNMGARKNAKFLPL 240  
DB 197 RLHFTKEQNNPNRSLKPIPTOLLGRTHATGLRKVVRELFTNMGARKNAKFLPL 256  
QY 241 TGEKFGDPLGYEDVPIELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVFQAN 300  
DB 257 TGEKFGDPLGYEDVPIELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVFQAN 316  
QY 301 NPEALKTQNLREKIFALBGTGTGSSFEHMSOEGFSAITNSGILLSTVGSVDWAG 360  
DB 317 NPEALKTQNLREKIFALBGTGTGSSFEHMSOEGFSAITNSGILLSTVGSVDWAG 376  
QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAAILLRNRVQSLVGLGAPRYQHIGLVAMER 420  
DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAAILLRNRVQSLVGLGAPRYQHIGLVAMER 436  
QY 421 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYYEQTRGGQVSVCP 480

DB 437 QNTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYYEQTRGGQVSVCP 496  
QY 481 PRGQARWQCDVLYGEOGQPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
DB 497 PRGQARWQCDVLYGEOGQPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSORLIAGSKLSPLOYFGQSLSGGQDLTMDGLVDTLVGAOHHVLLRSQ 600  
DB 557 HGTSGSGISPSHSORLIAGSKLSPLOYFGQSLSGGQDLTMDGLVDTLVGAOHHVLLRSQ 616  
QY 601 PVLAVKAIMFNPFRVARNVFECDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVYT 660  
DB 617 PVLAVKAIMFNPFRVARNVFECDQVYVKGKAGEVRVCLHVQKSTRDLREGQIQSVYT 676  
QY 661 YDLALDSGRPHSAVFNETKSTRQTVLGLTQCTETIKLQPNCTEDPVSPIVLRNLP 720  
DB 677 YDLALDSGRPHSAVFNETKSTRQTVLGLTQCTETIKLQPNCTEDPVSPIVLRNLP 736  
QY 721 SLVGTPLSAFGLNPLVLAEDAQRLLFTALFPPKKNCGNDNICQDDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFGLNPLVLAEDAQRLLFTALFPPKKNCGNDNICQDDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
DB 797 GPREFNVTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPEENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900  
DB 857 SGALKSTSCSINHPIPEENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 916  
QY 901 QLELPVKYAVYMWVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 960  
DB 917 QLELPVKYAVYMWVTSHGVSSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLSLVFLVPV 976  
QY 961 RLNOTVIWDRPQVTFSENLSSCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDIP 1020  
DB 977 RLNOTVIWDRPQVTFSENLSSCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDDIP 1036  
QY 1021 FPGIOEFPNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEP 1080  
DB 1037 FPGIOEFPNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSQTEP 1096  
QY 1081 KVEPPEVNPPLPLIVGSSVGGLLALLIITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137  
DB 1097 KVEPPEVNPPLPLIVGSSVGGLLALLIITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153  
RESULT 8  
US-09-902-481A-5  
; Sequence 5, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-5

Query Match 99.3%; Score 5836; DB 10; Length 1137;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60  
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 61 RLOVPVEAVNMSLGLSLAATTPPQLLAGCPTVHOTCSENTYVKGCLFLFGSLNRQQPQK 120  
DB 61 RLOVPVEAVNMSLGLSLAATTPPQLLAGCPTVHOTCSENTYVKGCLFLFGSLNRQQPQK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSSEF 180  
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSSEF 180

QY 181 RIHFTFKFQNNPNRSLIKPTTOLLGRTHATGLKVVRELENTNGARKNAFKILFL 240  
DB 181 RIHFTFKFQNNPNRSLIKPTTOLLGRTHATGLKVVRELENTNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAPRSEKSRQELNTVASKPDRDHVFOAN 300  
DB 241 TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAPRSEKSRQELNTVASKPDRDHVFOAN 300

QY 301 NFEALKTQONQUREKIIPAEIGTQSGSSSFEHMSQEGFSAATISNGPLLSTVGYSDWAG 360  
DB 301 NFEALKTQONQUREKIIPAEIGTQSGSSSFEHMSQEGFSAATISNGPLLSTVGYSDWAG 360

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVILGAPRYOHIGLVAMFR 420  
DB 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVILGAPRYOHIGLVAMFR 420

QY 421 QNTGAWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYQOTGGQSVSCPL 480  
DB 421 QNTGAWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYQOTGGQSVSCPL 480

QY 481 PRGORARWQCDAVLYGEOQPGWRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILP 540  
DB 481 PRGORARWQCDAVLYGEOQPGWRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILP 540

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLCYFGQSLGGQDLTMDGLVLTGAGQCHVILLRSQ 600  
DB 541 HGTSGSGISPSHSQRIAGSKLSPRLCYFGQSLGGQDLTMDGLVLTGAGQCHVILLRSQ 600

QY 601 PVLRVKAIIMEFNPREVARNVFCNDQVXGKAGEVRVCLHVQKSTRDLRREGOIQSVVT 660  
DB 601 PVLRVKAIIMEFNPREVARNVFCNDQVXGKAGEVRVCLHVQKSTRDLRREGOIQSVVT 660

QY 661 YDLALDSGRPHSRVAVNETKNSRTRQTQVLGTQCTETLKLQLPNCIEDPVSPVILRLNF 720  
DB 661 YDLALDSGRPHSRVAVNETKNSRTRQTQVLGTQCTETLKLQLPNCIEDPVSPVILRLNF 720

QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFTALPFPFKNCGNDNICQDDLSITPFSMSLDCLVYG 780  
DB 721 SLVGTPLSAFGLNRPVLAEDAQRLEFTALPFPFKNCGNDNICQDDLSITPFSMSLDCLVYG 780

QY 781 GREFNWTVTVRNDGDSYRTQVTFPPFLDLSYRKUSTLQNSORSRWELACESASTEV 840  
DB 781 GREFNWTVTVRNDGDSYRTQVTFPPFLDLSYRKUSTLQNSORSRWELACESASTEV 840

QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
DB 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900

QY 901 QLELPVKYAVMYVTSHGVSSTKYLNTASENTSRVWQHGYQVSNIGORSILPISLFLVPV 960  
DB 901 QLELPVKYAVMYVTSHGVSSTKYLNTASENTSRVWQHGYQVSNIGORSILPISLFLVPV 960

QY 961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSSDFLAELKAPVWNCSTAVCORIQCDIP 1020  
DB 961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSSDFLAELKAPVWNCSTAVCORIQCDIP 1020

QY 1021 PFGIOEBFNATLKGNSLFDWYIKTNSHLLIVSTAEILFNDVSFTLLPFGQCAFVRSQTEF 1080

DB 1021 PFGIOEBFNATLKGNSLFDWYIKTNSHLLIVSTAEILFNDVSFTLLPFGQCAFVRSQTEF 1080  
QY 1081 KVEPPEVPNPPLPIVGGSVGGLLALLTAALYKLGPFKROYKMMSEGGPPGABPQ 1137  
DB 1081 KVEPPEVPNPPLPIVGGSVGGLLALLTAALYKLGPFKROYKMMSEGGPPGABPQ 1137

RESULT 9  
US-09-902-481A-6  
; Sequence 6, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 6  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-6

Query Match 99.2%; Score 5832; DB 10; Length 1137;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60  
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 61 RLOVPVEAVNMSLGLSLAATTPPQLLAGCPTVHOTCSENTYVKGCLFLFGSLNRQQPQK 120  
DB 61 RLOVPVEAVNMSLGLSLAATTPPQLLAGCPTVHOTCSENTYVKGCLFLFGSLNRQQPQK 120

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSSEF 180  
DB 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSSEF 180

QY 181 RIHFTFKFQNNPNRSLIKPTTOLLGRTHATGLKVVRELENTNGARKNAFKILFL 240  
DB 181 RIHFTFKFQNNPNRSLIKPTTOLLGRTHATGLKVVRELENTNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAPRSEKSRQELNTVASKPDRDHVFOAN 300  
DB 241 TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAPRSEKSRQELNTVASKPDRDHVFOAN 300

QY 301 NFEALKTQONQUREKIIPAEIGTQSGSSSFEHMSQEGFSAATISNGPLLSTVGYSDWAG 360  
DB 301 NFEALKTQONQUREKIIPAEIGTQSGSSSFEHMSQEGFSAATISNGPLLSTVGYSDWAG 360

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVILGAPRYOHIGLVAMFR 420  
DB 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVILGAPRYOHIGLVAMFR 420

QY 421 QNTGAWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYQOTGGQSVSCPL 480  
DB 421 QNTGAWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYQOTGGQSVSCPL 480

QY 481 PRGORARWQCDAVLYGEOQPGWRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILP 540  
DB 481 PRGORARWQCDAVLYGEOQPGWRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVILP 540

```
QY 541 HGTSGSISPSHSORISAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGAGQGHVLLRSQ 600
DB 541 HGTSGSISPSHSORISAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGAGQGHVLLRSQ 600
QY 601 PVLKVKAIMFNPREVARNVFCNDQVVKGKEAGEVCLHVQKSTRDLREQIQSVVT 660
DB 601 PVLKVKAIMFNPREVARNVFCNDQVVKGKEAGEVCLHVQKSTRDLREQIQSVVT 660
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVGLGLTQTCETKLQLPNCIEDPVSPIVLKLPF 720
DB 661 YDLALDSGRPHSAVNETKNSRTRQTVGLGLTQTCETKLQLPNCIEDPVSPIVLKLPF 720
QY 721 SLVGTPLSAFAGNLRPVLAEDAQRFTALPPFKKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFAGNLRPVLAEDAQRFTALPPFKKNCNDNI CODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
QY 841 SGALKSTSCSINHPIPPENSESVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEP 900
DB 841 SGALKSTSCSINHPIPPENSESVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEP 900
QY 901 QLELPVKYAVYVMTSHGVSTKYLNTFASNTSRVMQHOYQVSNLQORSPLISLVLVPV 960
DB 901 QLELPVKYAVYVMTSHGVSTKYLNTFASNTSRVMQHOYQVSNLQORSPLISLVLVPV 960
QY 961 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIQCIP 1020
DB 961 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIQCIP 1020
QY 1021 FFGIQEEFNATLKGNLSPDVIYKTSNHLIIIVSTABILFNDVSFTLLPGQGFVRSQTET 1080
DB 1021 FFGIQEEFNATLKGNLSPDVIYKTSNHLIIIVSTABILFNDVSFTLLPGQGFVRSQTET 1080
QY 1081 KVEPFEVNPPLIIVGSSVGLLALITAALYKLGFFKQYKDMMSGGPPGABEQ 1137
DB 1081 KVEPFEVNPPLIIVGSSVGLLALITAALYKLGFFKQYKDMMSGGPPGABEQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimooka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4

Query Match 99.1%; Score 5823.5; DB 9; Length 1152;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1123; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 FNLDTENAMTQENARFGQSVVOLQGSRRVWVAGPQBIIVAAVNRGSLYQCDYSTGSCPTI 60
DB 17 FNLDTENAMTQENARFGQSVVOLQGSRRVWVAGPQBIIVAAVNRGSLYQCDYSTGSCPTI 76
QY 61 RLQVFEVAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 120
```

RESULT 11



138 LPVSRQECRQEQDIFVLIDGSGISIRNFATMMFVRAVISQFQSPFQSLMFSNKF 197  
181 RIHFTKEFONNPNRSLKIPITOLLGRTHATGLKRVRELFNITNGARKNAKILFLL 240  
198 QTHFTFEERTSNPLSLASVHQLQGYTATAIQNVVHERFHASVGARRDAIKLIVI 257  
241 TQCEKGDPLGYEDVPELDRGVIVYVLGFGDAFRSEKSRHSLNTVASKPRDHVFOAN 300  
258 TDCKKSGDSLDYKDVIPMDAAGIIRYALGVGLAFQNRNSWELNDIASKPSQEHFKVE 317  
301 NFALKTVONOLREKIPALIEGTOTGSSSFEHMSQEGFSAATISNGPLISTVGSVDWAG 360  
318 DFDALDKDIONOLKEKIPALIEGTETISSSFELEMAQEGFSAVTPDGVLCAGSFTWSG 377  
361 GVFLYTSKEKSTPINTRVDSQNDMDAYLGYAAAILRNVRQSVLVLGAPRYOHIGLVAMPR 420  
378 GAFLYPPNNSPTFINMSQENVDNRDSYLGYSTELALWKGVQSVLVLGAPRYOHIGKAVIFI 437  
421 QNTGMWESNANVKGTQIGAYFCASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQSVVCP 480  
438 QVSRQWRKMAKVEIGTQIGSYFCASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQSVVCP 497  
481 PRGQBARQCDANLYXGEOQWGPFGAALTVDLVGNGDKLTDVAIGAPCEDNRCAVILP 540  
498 PRGWR-RWMCDAVLYGEGQHPWGPFGAALTVDLVGNGDKLTDVVGAPGEENRGAVILP 556  
541 HGTSGSGISPSHSORLQAGSLSPRLQYFGQSISGQDLTMDGLVLTGACQHVLLRSQ 600  
557 HGVLPSPISPSHSORLQAGSLSPRLQYFGQALSGQDLTQDGLVDLAVGARQVILLRTR 616  
601 PVLRYKALMEFNPREVARNFECDQVWKGAESEVRVCLAVOKSTRDLREGOQSUVT 660  
617 PVLWYGVSMQFIPAEIPRSAFECEQVVSQVSEQLVQSNICLYIDKRSKLLGSRDLQSSVT 676  
661 YDLALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETKLQLPNCIEDPVSPIVLRNF 720  
677 LDALAPGRLSPRALFQETKRSLSRVRVLGKAHCENFNLLFSCVSDSVIPIILRNF 736  
721 SLVGTPLSPAGNLRPLVAEDAEQLFTALPPFEKNCNDNICQDLSITFSMSLDCLVYG 780  
737 TLVGKPLAFNLRPLMALAARYFTASUPFEKNCADHICQDNLGISFSFGLKSLVYG 796  
781 GPREFNVTVVRNDGSDSVYQTFPPDLDSVRKYSTLQNRQSQRWRLACESASTEV 840  
797 SNLELAEVVMWMDGSDSYCTITTFSPAGLSRYVYABQKQQLRSLHLC--CSAPVG 854  
841 SGALKSTSCSINHPIPENSEVTFNTFDVDSKASLGNKLLKARVTSNNMPTWKTET 900  
855 SQGTWSTSCINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANVSSENNIPRSTKIP 914  
901 QLELPKYAYVMVTSHGYSKYLANFTAS-ENTSRVMQHOYQVSNIGORSPLISVFLVP 959  
915 QLELPKYAYIVVSSHEQTKYLNFSSESEKESHVAMHRYQNNILGORDLPSVINFVVP 974  
960 VRLNQTIVNDRPOVTFSENLSSTCHTKERLPSHSDFLAEIRKAPVNVCSIAVOCIOQDI 1019  
975 VELNQAVMDVSHVQNPDLSCSEKIAIPASDFLAHQKQVLDLSCIAGLRFRCDV 1034  
1020 PPFQIQEFNATLKNLSFDWYIKTSHNHLIIVSTAEIIFNDSVFTLLPQCGAFVRSOTE 1079  
1035 PSFSVQDELDFTLKNLSFGWVQILQKKVSVVVAEIIIFDTSVYSQLPQCAFMAQTI 1094  
1080 TKVEPEVFNPLIVGSSVGGILLIALLIATAALYKLGPFKQYKQWMBE 1128  
1095 TVLEKYKVRNPIPLIVGSSVGGILLIALLIATAVLYKVGPFKQYKQWMBE 1143

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US2003007728A1

; GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: NO. US2003007728A1el Human 2  
FILE REFERENCE: 27866/35004  
CURRENT APPLICATION NUMBER: US/09/891,943  
CURRENT FILING DATE: 2001-06-26  
PRIOR APPLICATION NUMBER: 09/193,043  
PRIOR FILING DATE: 1998-11-16  
PRIOR APPLICATION NUMBER: 08/286,889  
PRIOR FILING DATE: 1994-08-05  
PRIOR APPLICATION NUMBER: 08/362,652  
PRIOR FILING DATE: 1994-12-21  
PRIOR APPLICATION NUMBER: 08/943,363  
PRIOR FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 4  
LENGTH: 1163  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-891-943-4

Query Match 58.2%; Score 3423; DB 10; Length 1163;  
Best Local Similarity 60.2%; Pred. No. 2.8e-307;  
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVVGAPQEIIVAAQRGSLYQCDYSTGSCBPI 60  
DB 20 FNLDTEELTAFRVDSAGFGDSVVQYANSVVVVGAPQKIIAANOIGELQCCGYSTGACBPI 79  
QY 61 RLQVPEAVNNSLGLSLAATSPOLLACGPTVTCSTENTYVKGCLCFGLFSGNLBQQOK 120  
DB 80 GLQVPEAVNNSLGLSLAATSPOLLACGPTVTCSTENTYVKGCLCFGLFSGNLBQQOK 137  
QY 121 FPEALRGCPQSDITAFILDCSGSIIIPHDPRMKELVSTIMEQLKSKTFLSLMYSBEF 180  
DB 138 LPVSRQECRQEQDIFVLIDGSGISIRNFATMMFVRAVISQFQSPFQSLMFSNKF 197  
QY 181 RIHFTKEFONNPNRSLKIPITOLLGRTHATGLKRVRELFNITNGARKNAKILFLL 240  
DB 198 QTHFTFEERTSNPLSLASVHQLQGYTATAIQNVVHERFHASVGARRDAIKLIVI 257  
QY 241 TDCEKGDPLGYEDVPELDRGVIVYVLGFGDAFRSEKSRHSLNTVASKPRDHVFOAN 300  
DB 258 TDCKKSGDSLDYKDVIPMDAAGIIRYALGVGLAFQNRNSWELNDIASKPSQEHFKVE 317  
QY 301 NFALKTVONOLREKIPALIEGTOTGSSSFEHMSQEGFSAATISNGPLISTVGSVDWAG 360  
DB 318 DFDALDKDIONOLKEKIPALIEGTETISSSFELEMAQEGFSAVTPDGVLCAGSFTWSG 377  
QY 361 GVFLYTSKEKSTPINTRVDSQNDMDAYLGYAAAILRNVRQSVLVLGAPRYOHIGLVAMPR 420  
DB 378 GAFLYPPNNSPTFINMSQENVDNRDSYLGYSTELALWKGVQSVLVLGAPRYOHIGKAVIFI 437  
QY 421 QNTGMWESNANVKGTQIGAYFCASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQSVVCP 480  
DB 438 QVSRQWRKMAKVEIGTQIGSYFCASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQSVVCP 497  
QY 481 PRGQBARQCDANLYXGEOQWGPFGAALTVDLVGNGDKLTDVAIGAPCEDNRCAVILP 540  
DB 498 PRGWR-RWMCDAVLYGEGQHPWGPFGAALTVDLVGNGDKLTDVVGAPGEENRGAVILP 556  
QY 541 HGTSGSGISPSHSORLQAGSLSPRLQYFGQSISGQDLTMDGLVLTGACQHVLLRSQ 600  
DB 557 HGVLPSPISPSHSORLQAGSLSPRLQYFGQALSGQDLTQDGLVDLAVGARQVILLRTR 616  
QY 601 PVLRYKALMEFNPREVARNFECDQVWKGAESEVRVCLAVOKSTRDLREGOQSUVT 660  
DB 617 PVLWYGVSMQFIPAEIPRSAFECEQVVSQVSEQLVQSNICLYIDKRSKLLGSRDLQSSVT 676  
QY 661 YDLALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETKLQLPNCIEDPVSPIVLRNF 720  
DB 677 LDALAPGRLSPRALFQETKRSLSRVRVLGKAHCENFNLLFSCVSDSVIPIILRNF 736

```

Db 471 TCSLSVAGAPQYKHG--AVPELOKGEASFLPVLEGEQMSYFGSELCPVDIMDGS 528
Qy 472 TDVLVIGAPHYTEQTGGQVSCPIPRGQARWQCDVLYGEOGPWGRFGNALTVLGDV 531
Db 529 TDFLLVAAPFYHVHGEGRVYVYRUSE-QDGSFLARILSGHGFNTNARFGFMAWMDL 587
Qy 532 NGDKLTDVAIGAP---GEEDNR--GAVLFGTSGSGISPSHSQRIAGSKLSPRLQYFG 585
Db 598 SQDKLTDVAIGAPLEPGFGADGASFGSVIYNG-HWDGLSAPSQRIRASTVAPGLQYFG 646
Qy 596 QSLSGGDLTMDGLVDLVGAGCHVLLRSQVLRKALMEPNPREVARNVFCNDQVVK 645
Db 647 MSWAGGFDSIGDGLADITVTLGQAVFRSREVRLKVSMAFTPSALP-----IG 696
Qy 646 GKEAGEVRVCLHVQKSTRDLREGIQSVVYDLDGSRPHRAVFNETKNSRTRQVQ 705
Db 697 ENGVMVNRVLCPEI-SSVITASSGLREALNFTLDVQKRRRLQCSDEVSCGLGLEM 755
Qy 706 LGLTQCTETLKLQLEN-----CIEDPVSPVLRLNLSVGLTSLASRNLRLPVLAEDARLP 761
Db 756 SSGSQLCEDL-LLMTEGELCEEDCFNSAKVSYQL-QTEGQTDHPQPILDRYTEPFA 813
Qy 762 TALPFPEKNCNDNICQDDLSITFSFMSLDCLVWGPREFVTVTRNDGDSYRTOVTF 821
Db 814 IFOLPYEKACQKLCFVAELQLA-TVSQQLVVLGLTKELTALINLTNSGDSYVTSNAL 872
Qy 822 FPLDLSTYKNTLQNRQSRWRACASASTEVSGALKSTSCSINHPIPPENSEVTFN 881
Db 873 NYPRNLQ-----LKRMPKPPFNQCDPQPV---ASVLMNCRIGHFVL-KRSSAHVS 922
Qy 882 ITFVDVSKASLGNKLLKANVTSEN---MPTNKTETFO---LPLPKYAVVYVVTSHG 934
Db 923 VWQLEENAFNRTADIITVTNSNERSLANETHLQFRGHFVAVLSKPSIMYVNTQGG 982
Qy 935 VS--TKYLNFTASENTSRVMQHQYQVSNLQSRSLPISLVLVPLVNLQTVIWRPQVTPS 992
Db 983 LSHKKEFLFVHGEN---LFGAEYQ-----LQICVPTKRLQVAAVKKLRT 1027
Qy 993 ENLSSTCHTKELPSHSDLEALRKAPVYVNCISVACQRIQCDI PFPGIQEFNATLKGNL 1052
Db 1028 QASTVCTVTSQERACAYSS-VQHVSEWHSVSCVIA-----SDKENVTVAAEI 1072
Qy 1053 SPDWYIKTSHNHLIVST-----AEILFNDVSFTLLPQCAFVSQETKVEPPEVNP 1106
Db 1073 SWD-----HSELLKDVTELQILGEISFNKSLYGLNAENH--RTKITVFLVLDKHYH 1124
Qy 1107 LPLVGVSVGGLLIALALTAALYKLGPKRQYKMMSE 1144
Db 1125 LPIIIKSVGGLLVILVILVILFKCGFFKRYQQLNLE 1162

```

## RESULT 9

```

ID ITA1_HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R.; Epstein M.R.; Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
RT 1 subunit."
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-

```

```

CC CC E-R IN COLLAGEN.
CC CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC CC ASSOCIATES WITH BETA-1.
CC CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC CC -!- DOMAIN: THE INTEGRIN 1-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC CC WITH 1-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC CC -!- SIMILARITY: Contains 1 VWFA domain.
CC CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC CC -!- DATABASE: NCBIPROV: NCBIPROV: NCBIPROV: NCBIPROV: NCBIPROV: NCBIPROV:
CC CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".
DR PIR: A45226; A45226.
DR PDB: 1QCS; 17-MAY-00.
DR Genew: HGNC:6134; ITGA1.
DR MIN: 192968; -.
DR GO: GO:0008305; C:integrin complex; TAS.
DR GO: GO:0004895; P:cell adhesion receptor activity; NAS.
DR GO: GO:0005518; P:collagen binding; TAS.
DR GO: GO:0007160; P:cell-matrix adhesion; NAS.
DR InterPro: IPR000413; Integrin_alpha.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.
DR SMART: SMART; SMO0191; Int_alpha; 5.
DR SMART: SMART; SMO0327; VWA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS0234; VWFA; 1.
DR Integrin: Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure.
FT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1114 1136 POTENTIAL.
FT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT 17 75 FG-GAP 2.
FT DOMAIN 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GEFK MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC... (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC... (POTENTIAL).

```



FT DISULFID 998 1023 BY SIMILARITY.  
FT DISULFID 1031 1047 BY SIMILARITY.  
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 773 773 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 823 823 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 846 846 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 1167 AA; 128983 MW; B83331L15DCCCFD CRC64;  
Query Match 19.6%; Score 1165.5; DB 1; Length 1167;  
Best Local Similarity 29.1%; Pred. No. 1.8e-69;  
Matches 362; Conservative 213; Mismatches 465; Indels 205; Gaps 43;  
QY 5 VLLLTALITLCHGFNLDTEA--MTQENARGFGQSVQLOGSRVVGAPQELVAANORGS 62  
DB 8 LLCMASLKPQGFANMDVMAWTALQPGAPAVLSLLHLDPEN-----NOTCLLVARSS 62  
QY 63 -----LVQDSTGSCPIRLQVVEAVNMSLGLSLAATT--SPFQLLAC--GPTVHQT 113  
DB 63 NENTAALYRCALSI-SPDEIAQ-PVEHICMPKRGYQVTLVGNHGVLCVQVQAKFR 120  
QY 114 SENTVVGKLCFTFGNLRQOPKPEALRG-----C----- 144  
DB 121 SUNSELTGACSLTFLNLDLQQAIFYSDLEGFLDPGARHVDGDCYSGKSGTGEKKSARR 180  
QY 145 -----PODSDIAFLIDGSGIIPHDFRMKEFVSTVMEQL--KKSKTFLGIMQVS 193  
DB 181 RRTVBEDEEINGTEIAVLVDGSGISGSDPDQKKNFTSTMRNFEKFCFECFALVQVG 240  
QY 194 EEFRIHFTKPEFQNNPNSLVKPTTOLLGRTHATGINKVRELFNITNGARKNAFKIL 253  
DB 241 AVIQTEFDLQESRDINASLAKVQSVQVKEVTKTASAMQHVLDNITFIPSGRKKALKVM 300  
QY 254 VVITGEKGDPLGVEDVTPEDRQGVIRYVVGDPAPSEKSRQBLATIASKPPRDVFE 313  
DB 301 VVITDGDIFGDPENLTVTINSPEMQGVRFALQVGRFANNYRELKLLIASDPKEATFP 360  
QY 314 QVNNFEALKTIONLREKIFAJEGTQSGSSSFHEMGEQFSAAITNSGP--LLSTVGSY 372  
DB 361 KVTNYSAIDGLLSKLQORIVHMEGT---VGDALQYLAQTGPSAQILDKGVLLGTVGAF 417  
QY 373 DWAGGVFLY-TSKERSTENMT-RVDS-DMNDAYLGAAAILNRVQSLVLAGAPRYQHI 429  
DB 418 NMSGGALLYSTONGRCFLNATAKEDSRVQSYLSGLSLAVLHKAGHSYVAGAPRHKUR 477  
QY 430 GLVAFRQNTGHWESNA---NVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYRQT 486  
DB 478 GAVFELRKEDR--BEDAFVRRIEGEQMGSYFGSVLCPVDIDWDGTTDFLLVAAPFYHIG 535  
QY 487 RCGQSVVCLPFGQARWOCDAVLGEQCPWGRFGAALTVLGDVNGDKLTDAVAGAP-- 544  
DB 536 BEGRVYVQVPE-QDASFSLAHLTSGHPGLTNSRFGFAMAAVGDINDQKFTDVAIGAPLE 594  
QY 545 ----GREDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGDLTMDGLV 600  
DB 595 GFGAGDGYSGVYIYNGHSG-GLYDVSFQQRASSVASGLHYFGMSVSGGLDFNGDGLA 653  
QY 601 DLTVGAQGHVLLIRSQPLRVKAIYFNPREVARNVFCNDQVYVKGAGEVRVCLHVQK 660  
DB 654 DITVGRSDSAVVLRSRPPVLDLTVSMFTFP-----DALEPMVFIKGM--DYNLCFEVDS 703

QY 661 S---TRDLREGQIQSVVTFYDLALDSCRPSPRAVFNETKNSTRQTQVLGLATYC----- 712  
DB 704 SVVASBEGLEAMFLNFTVDVV-----TKQORLQCEDSSGCGCLRKWN 748  
QY 713 -----ETLKLOLENCHIEDPVSPIVRLNLSLVGTPLSAPGNLR---PVLAE 756  
DB 749 GGSFLCBFWLSTEEEL-----CEDCFENITIKVYIS-----PQTSGRRDYFNPTU--D 797  
QY 757 AQRLLTALF--PPEKNGNDNIQDDLSITFSFMSLDCLVVGGPREFNVTVTVNRNDEGS 814  
DB 798 HYKEPSAIFQLPYEKDCKNVFCIAEIQLTN-ISOQELVVGTVKEVTMNLSTNSGDS 856  
QY 815 YRTQVTFPFLDLSYRKVSTLQNSORSORSLACESASSTEVSGALKSTSCSINHPIFE 874  
DB 857 YTMNMAIYFNQLQFKKI-----QKEVSPDVQCDPKPV---ASVLWMNCKIGHPL-K 906  
QY 875 NSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEFQLELPKVAVMVTVSHG 934  
DB 907 RSSVNVSVTWQLESVFENETADITVTSNKSLEARETR---SLQFRAFIATLSR-- 961  
QY 935 VSTKYLNTASENTSRVHQYQVSNLQORSPLISLVFLVPLNQTIVWDRPOVTFSEN 994  
DB 962 PSVMYAN--TSQSPSDHKEFFPNVHGENLFGAVFQLQICVPIKLQDF-----QIVRVKN 1013  
QY 995 LSST-----CHIKERLPSHDSFLAELRKAPVWNCIAVCQRIQCDIPFFGIOEENATLK 1049  
DB 1014 LTKQDTHETQSGEPACGSDPVQHVKEHWSVVCAL-----TSNK 1053  
QY 1050 GNLSFDWIKYKSNHLLIVSTA-----EILFNDSVFTLLPGQAFVRSOTETKVEFF-- 1101  
DB 1054 ENVTVAEISVGHTKQLLRDVSELPIGELISFNKSLYEGLNAE-----NHRKTIWIFL 1107  
QY 1102 --EVPNPLPLVSGVGLALLALITAAALYKLGFFKRYKDMSE 1144  
DB 1108 KEETRSPLIIGSSIGLLVWVILALFCGFFKRYKQENLE 1152  
RESULT 8  
ITAE HUMAN  
ID ITAE HUMAN STANDARD; PRT; 1179 AA.  
AC P38570; QNZU9;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1  
DE antigen) (CD103 antigen) (Integrin alpha-IEL).  
GN ITGAE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.  
RC TISSUE=Leukemia, and Lymphocytes;  
EX MEDLINE=94164962; PubMed=8119947;  
RA Shaw S.K.; Cepek K.B.; Murphy E.A.; Russell G.J.; Brenner M.B.,  
RA Parker C.M.;  
RT "Molecular Cloning of the human mucosal lymphocyte integrin alpha E  
RT subunit. Unusual structure and restricted RNA distribution.",  
RL J. Biol. Chem. 269:6016-6025(1994).  
RN [2]  
RP REVISIONS TO 88-114.  
RA Parker C.M.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 53-1179 FROM N.A.  
RC TISSUE=Fetal kidney;  
EX MEDLINE=20138496; PubMed=10673275;  
RA Touchman J.W.; Anikster Y.; Dietrich N.L.; Maduro V.V.; McDowell G.,  
RA Shotelersuk V.; Bouffard G.G.; Beckstrom-Sternberg S.M.; Gahl W.A.,  
RA Green B.D.;  
RT "The genomic region encompassing the nephropathic cystinosis gene  
RT (CTNS): complete sequencing of a 200-kb segment and discovery of a

```
Db 342 SFNKLSSGSGADLSKSHAVGVAGKAWAGGFLDLREDLQAGTFVQCBPLTSDVRGGY 401
Qy 404 LGYAAA-IILRNRYQSLVGLGAPRYOHICLVAMFR--ONTGWESNANVKGFOICAYGAS 460
Db 402 LGYVAMWTSRSSRPLLAAGAPRYOHVGVLLFOAPAGGRWNTQIKGFOICSGYFGE 461
Qy 461 LCSVDVDSNGSTDLVLGAPHYYQTRGGQVSVCLPRGQRARWCDVAVYGEQGPWGR 520
Db 462 LCSVDLDQDGRABLLIGAPLFFGQGRGRVFTY---QRRSLFEMVSELQDGPYPLGR 518
Qy 521 FGAALTIVLDVNGDKLTDVATGAPGRNRCNAVLFECTSGSISPSHSORLAGSKLSR 580
Db 519 FGAALTITDINGDLTDVAGAPLEB--QNAVTFNGKXP-GUSPQPSORIOGAQVFP 575
Qy 581 LOYFQSLSGQDUTMDGLVDLTVGAGHVLNLSQPVLRVKAIMEFNPREVARNVFECN 640
Db 576 IRWEGRSIHGKDLGGDLADVVVGAEGRVVLSRVVDVWTELSFSPEEIPVHEVECS 635
Qy 641 DOVVKGEAG-EVAVCLFVQKSTDRDLREGQIOSVVTYDLADSGRPHSRVAVNETKST 699
Db 636 YSAREEQHGKGLKACFRIRKPLTPQ--FQGLLANLSTYQLDGHMRSRGLPFGSHEL 693
Qy 700 RROTQVLGLTQCTETLKLQPNCTEDPVSPIVLRLNFSLV---GTPLSAPFN-LRPVLAE 755
Db 694 SGNSTIIP-DKSLDFHFPICIQDLISPINVSLNFSLEEBCTPRDQKRAMQPILRP 752
Qy 756 DAQRLFTALPFFKNCNDNICQDDLITFMSLDCLVGGP-----REFNVTVTVRN 809
Db 753 SIHTV-TKEIPFKNCEGCKCEANLTLSPPARS-----GPLRLMSSASLAVETLSN 804
Qy 810 DGEDSYRTVTFPPDLVSRKSVSTLQVRSQSRWRLACESASSTEVSGAL-KSTSCIN 868
Db 805 SGEDAVYWRDLDFRGLSFRKVMQLQ---PHSRMPVSCHEL--TEGSSLLTKLKNSV 859
Qy 869 HPIPPENSEVNTITFDVDSKASIGNKLLKANVTSEN-NMPTNKTEFOLFLPKYAVY 927
Db 860 SPIFKAQEVSQVWNTLNNSEWDFVLNGTVHCENENSSLQEDNSAATHIPVLXPVN 919
Qy 928 MVTSHGVSTKYNLFTASENSTRVQHOXOVNLSQESLPSLVPLVPLVNLQTVIWRP 987
Db 920 ILTKEQENSTLYISFTPKGKTQVQVHYQV----RIQSAVDHNMPT-LEALVGVPRP 973
Qy 988 Q-----VTFSENLS---TCHTKE-RLPSSHDFLAEALRKA PVNCSIAVQCQICDIPFF 1038
Db 974 HSEDLITYTWSQTDPLVTVCHSEDLKRPSE---AEQFCILPV-----QFRCPIVE- 1021
Qy 1039 GIQEEFNATLKNLSFDWYIKTSHHLLIVSTAEILEPNDVSFTLLPQOGAFVRSOTETKV 1098
Db 1022 --RWEILIQTVGFELSKIRAS-STLSLCSLSVSFNSKHFHYLXGSKA-SEAQVLVKV 1077
Qy 1099 EPFEVNPPLPLVGVSSVGLLALLALITAAALYKLGFFKRYQKDMW-SBGSGPPGAEP 1152
Db 1078 DLIHEKEMLVHVLVLSGIGGLVLLFLIALYKVGFFKRNKRYKHEADGGVPNGSP 1132

RESULT 7
ITAE MOUSE STANDARD; PRT; 1167 AA.
AC Q60677;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, last annotation update)
DI Integrin alpha-E precursor (Integrin alpha M290).
GN ITGAE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AKR;
RX MEDLINE=95187992; PubMed=7882170;
```

```
RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,
RA Kilshaw P.J., Weis J.H.;
RT "Murine M290 integrin expression modulated by mast cell activation.";
RL Immunity 1:393-403 (1994).
CC -1- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOCYTE AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VWFA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; U12236; AAC52142.1; -.
CC HSSP; P11215; 1A8X.
CC MGD; MGI:1298377; Itgae.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 3.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS00234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Magnesium;
KW Calcium.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 1167 INTEGRIN ALPHA-E.
FT CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.
FT CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.
FT DOMAIN 20 1114 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1115 1137 POTENTIAL.
FT DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).
FT REPEAT ? ? FG-GAP 1.
FT REPEAT ? ? FG-GAP 2.
FT REPEAT 149 192 X-DOMAIN (EXTRA DOMAIN).
FT DOMAIN 193 384 VWFA.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 449 501 FG-GAP 4.
FT REPEAT 503 584 FG-GAP 5.
FT REPEAT 566 631 FG-GAP 6.
FT REPEAT 634 686 FG-GAP 7.
FT CA_BIND 514 522 POTENTIAL.
FT CA_BIND 578 586 POTENTIAL.
FT CA_BIND 646 654 POTENTIAL.
FT DOMAIN 185 191 GLU-RICH (ACIDIC).
FT SITE 1140 1144 GPFKR MOTIF.
FT DISULFID 72 83 BY SIMILARITY.
FT DISULFID 130 164 BY SIMILARITY.
FT DISULFID 698 754 BY SIMILARITY.
FT DISULFID 814 820 BY SIMILARITY.
FT DISULFID 884 898 BY SIMILARITY.
```

Db 1083 VYKQMLYLVLSIGGLILLILLIFVLYKVGFKRLKEMRAGRCVPNGIP 1136

RESULT 6  
ID ITAL\_MOUSE STANDARD; PRT; 1163 AA.  
AC P24063;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1  
alpha chain) (Leukocyte function associated molecule 1, alpha chain)  
(Cblaa).  
GN ITGAL OR LFA-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91268576; PubMed=2051027;  
RA Kaufmann Y., Tseng E., Springer T.A.;  
RT "Cloning of the murine lymphocyte function-associated molecule-1  
alpha-subunit and its expression in COS cells.";  
RL J. Immunol. 147:369-374 (1991).  
RN [2]  
RP SEQUENCE OF 24-42.  
RX MEDLINE=85188276; PubMed=3887182;  
RA Springer T.A., Teplow D.B., Dreyer W.J.;  
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion  
glycoproteins and unexpected relation to leukocyte interferon.";  
RL Nature 314:540-542 (1985).  
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,  
ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA  
INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL  
MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES  
AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L  
SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED  
LEUKOCYTES RECRUITMENT.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L  
ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: LEUKOCYTES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M60778; AAA39426.1; -.  
DR PIR; I56126; I56126.  
DR HSSP; P20701; 1LFA.  
DR MGD; MGI:96606; Itgal.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR02035; VWF A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

Signal; Magnesium; Calcium;

Repeat.  
KW SIGNAL 1 23  
FT CHAIN 24 1163  
FT DOMAIN 24 1084  
FT TRANSMEM 1085 1108  
FT DOMAIN 1109 1163  
FT REPEAT 39 88  
FT REPEAT 39 88  
FT REPEAT 148 334  
FT REPEAT 399 454  
FT REPEAT 455 514  
FT REPEAT 516 573  
FT REPEAT 576 628  
FT CB\_BIND 466 474  
FT CA\_BIND 528 536  
FT CA\_BIND 588 596  
FT SITE 1111 1115  
FT DISULFID 70 77  
FT DISULFID 108 126  
FT DISULFID 147 199  
FT DISULFID 651 703  
FT DISULFID 767 773  
FT DISULFID 840 856  
FT DISULFID 993 1009  
FT DISULFID 1017 1048  
FT CARBOHYD 86 86  
FT CARBOHYD 185 185  
FT CARBOHYD 270 270  
FT CARBOHYD 444 444  
FT CARBOHYD 668 668  
FT CARBOHYD 696 696  
FT CARBOHYD 724 724  
FT CARBOHYD 728 728  
FT CARBOHYD 776 776  
FT CARBOHYD 857 857  
FT CARBOHYD 880 880  
FT CARBOHYD 890 890  
FT CARBOHYD 899 899  
FT CARBOHYD 927 927  
FT CARBOHYD 1056 1056  
SQ SEQUENCE 1163 AA; 128343 MW; A7A3078489B8232F CRC64;

Query Match 26.0%; Score 1549; DB 1; Length 1163;  
Best local Similarity 34.2%; Pred. No. 5e-95;  
Matches 409; Conservative 216; Mismatches 464; Indels 106; Gaps 38;

Qy 1 MALRV---LLLTALT---CHGFNLDTENAMTFQENA-RGFQGVVQLGSRVWVGAPQ 52  
Db 1 MSFRIAGPRLLILGLQLPAKWSYNLDTRTQSFQAQGRHFGYVQLIEDG-VVVGAPG 59  
Qy 53 EIVAAHQSGLYOCYDSTGSCBPIRLQVPVEAVNMSGISLAATTPPQLLACGPTVHOT 112  
Db 60 E---GDNLTGLYHCRSTSSFCQPSVLSH-GSNHNTSKYLGWTLATDAKAGSLACDPLGSLRT 115  
Qy 113 CSNFFVYKGLCFGLFGSNLRQPOKPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKEF 172  
Db 116 CDQNTYLSGLCYLFPQSGEPMQLNPAYOECKKVDLVLFDSQSILDRKFEKILEP 175  
Qy 173 VSTVMEQLKSKTLFSLMAYSEBFRHFTPEKF-QNNPNPRSLVKPTQLLGRHTHTATGI 231  
Db 176 MKDVMKLSNTSYQAAVQFSTDCRTEFTLDYVKQKNDVLLGSVQPMFLTLNTPRAI 235  
Qy 232 RKVRELFNITNGARKNAFKILVITDGEKEGFGPLGYEDVIPLEADREG-----VIRYV 284  
Db 236 NYVAVHVFKEGSGARPDAITKVLVIITDG-----BASQKNTISAADITRYI 281  
Qy 285 IGVGDAFRSBSKROELNTIASKPPRDHVQVNNFEALKTIONQLREKIFAEGTQTGSSS 344  
Db 282 IGICKHFVSVQKQKTLHIFASEPVEBFVKLDTFEKLDKLTDLQRIYAEIGINRQDLT 341  
Qy 345 SFHEMSQEGFSAAITSGNPLLSTVGSYDAGGVF-LYTSKEKSTFINMTVEVSDMNDAY 403



```

Db 901 ENKASSKATQLELPKVIATYTWISROESTKTFNFATSEKOKKAEHRIRVNNLSQ 960
QY 964 RSLPISLVFLVRLNQTVINDRPQVPTSENLSSCTHKERLPDSDFLAELRKAPVYVC 1023
Db 961 RLALISINFPVLLNGVAVMDVMEAPQSL-PCVSEKPFQSHSDFLTQISRPFMLDC 1018
QY 1024 STAVCORIOCDPFGICGEEFNATLGNLSFDWYIKTSHNLLIYSTAEILLFNDSVFTLL 1083
Db 1019 STADCLQRCDFVPSVGEELDTLGNLSFGVRETLQKRVLVVSVARITFDTSVYSQL 1078
QY 1084 FQGAFAVSQETKVEPEVEPNFLPIYGVSSVGLLALITAAALYKLGFFRRQYKQMS 1143
Db 1079 FQGAFAVMAQEMVLEEDDEVYNAIPILMGSSVGLLALITATLYKLGFFRRHYKEMLE 1138
QY 1144 E 1144
Db 1139 D 1139

RESULT 5
ITAL HUMAN STANDARD; PRT; 1170 AA.
AC P20701; OX3746;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)
DE (CD11a).
GN ITGAL OR CD11A.
OS Homo sapiens (Human).
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RA MEDLINE=89139587; PubMed=2537322;
RA Larson R.S., Corbi A.L., Berman L., Springer T.;
RT "Primary structure of the leukocyte function-associated molecule-1
RT alpha subunit: an integrin with an embedded domain defining a protein
RT superfamily.";
RL J. Cell Biol. 108:703-712 (1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q.";
RL Genomics 60:295-308 (1999).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX MEDLINE=96036067; PubMed=7479767;
RA Qu A., Leahy D.J.;
RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha
RT L beta 2) integrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281 (1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX MEDLINE=96398682; PubMed=8805579;
RA Qu A., Leahy D.J.;
RT "The role of the divalent cation in the structure of the I domain
RT from the CD11a/CD18 integrin.";
RL Structure 4:931-942 (1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX MEDLINE=99425288; PubMed=10493852;
RA Kallen J., Welzbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA Cottens S., Weitz-Schmidt G., Hommel U.;
RT "Structural basis for LFA-1 inhibition uponlovastatin binding to the

```

```

RT CD11a I-domain";
RL J. Mol. Biol. 292:1-9 (1999).
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
CC ICAM3 AND ICAM4. IT IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA
CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL
CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES
CC AND MONOCYTES.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P20701-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P20701-2; Sequence=VSP_002738;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: LEUKOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWF A DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWF A domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAMB=PROW; NOTE=CD guide CD11a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00796; CAA68747.1; -
DR EMBL: AC002310; AAC31672.1; -
DR PIR: S03308; S03308.
DR PDB: 1LFA; 29-JAN-96.
DR PDB: 1ZON; 07-DEC-96.
DR PDB: 1ZOO; 07-DEC-96.
DR PDB: 1ZOP; 07-DEC-96.
DR PDB: 1COP; 07-AUG-00.
DR PDB: 1DSQ; 03-FEB-00.
DR PDB: 1MJN; 28-JAN-03.
DR PDB: 1MQ8; 14-JAN-03.
DR PDB: 1MQ9; 14-JAN-03.
DR PDB: 1MQA; 14-JAN-03.
DR Genew; HGNC:6148; ITGAL.
DR MTM; 153370; -
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0006928; P:cell motility; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00032; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; 3D-structure; Magnesium; Calcium; Repeat;
KW Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 1170
FT DOMAIN 26 1088
FT TRANSMEM 1089 1112
FT DOMAIN 1113 1170
FT REPEAT 92 91
FT REPEAT 149 349
FT DOMAIN 170 349

```



Db 8 LLLFTALATSLGMLDTEELTAFRVDSAGPDSVVOYANSVVVGAPOKITAANTQGLY 67  
QY 65 QDYSTGSCBPIRLOVPVEAVNMSLGLSLAATSPOLLACGPTVHOTCSNTYVKGLCF 124  
Db 68 QCGYSTGACEPIGLQVPPVAVNMSLGLSLASTTSPQLLACGPTVHHEOCRNMYLTGLCF 127  
QY 125 LFGSNLRQOPKEPPEALRGCPQEDSDLAFLIDGSGSIIPEDPRMKEFVSTWMEOLKSK 184  
Db 128 LLGPT--QLTORLPVSRQCPREQDVLFLIDGSGSISSENFAFMVFNANISQFQRS 185  
QY 185 TPLSMQYSEBFIHTTKEFQNNPNRSLVKEITQLLGRHTATGIRKVVRELFINING 244  
Db 186 TOPSLMQFNKEFQHTTPEFRSTNPLSLLASVHQLQFTVTATAIQNVHRELPHASVG 245  
QY 245 ARKNAKILVITDGEKPGDPLGYEDVPEADREGVIRYVIGVGDFAFSEKSOELNTIA 304  
Db 246 ARDATKILVITDGEKGSGLDYKDVIPYADAAGIIRYAIAGVGLAQQNWKELNDIA 305  
QY 305 SKPPRDHVQVNNFEALKTIQNLREKIPAEIGTQTGSSSSPEHEMSQSGFSAALTSNGP 364  
Db 306 SKPSQSHIFKVEDFDALKDIONOLKEKIPAEIGTETTSSTSSPELEMAQSGFSAVTPDGP 365  
QY 365 LLSTVGSYDAGVFLYTSKESSTFNMTRVDSMDNDVLYGAAAILLNRVQSLVIGAP 424  
Db 366 VLGAVGSFTWGAFLYPPNMPSTFNNMSQEWMDRDSYLGYSTELALWGVQSVLIGAP 425  
QY 425 RYCHIGLVAMFRQNTGMESNANVKTQICAYFGASLCSVDVDSNGSTDVLVIGAPHYVE 484  
Db 426 RYCHTGKAVIFQVSRQNMKAETVGTQIGSYFGASLCSVDVDTGSDTLVLVIGAPHYVE 485  
QY 485 QTRGGVSVCLPRGORARWCDVLYGQOPWGRFGAALTVLGVNGDKLTDVAIGAP 544  
Db 486 QTRGGVSVCLPRGWR--RWWCDVLYGQOPWGRFGAALTVLGVNGDKLTDVVIIGAP 544  
QY 545 GEEDNRGAVLYFHGTSGSGISPSHSQRTAGSKLSPRLQVFGOSLGGQDLTWDGLVDLTV 604  
Db 545 GEENRGAVLYFHGLVGPISPSHSQRTAGSLSSELQVFGALSGLGGQDLTQDGLVDLAV 604  
QY 605 GAQHVLLRSQVPLRVKAIMFNPREVARNVPEQNDQVQKKEAGEVRVCLVHVKSTRD 664  
Db 605 GARGQVLLRTPVLVGVSMQFIPABIPRSFAFECREQVWSQTLVQSNICLIDYDRSKY 664  
QY 665 RLREGIQSVVTVYDLALDSGRPHSAVFNETKSTRTQTVLGLTQTCETLKLQLPNCIE 724  
Db 665 LLGSRDLQSSVTLDLALDPGLSPRATQETKNSRLSVRVLGLKAKENFNLPLSPCVE 724  
QY 725 DVPSPVILRLNPSLYGTPLSAGNLRPVLAEDAQBLFTALPPEKXCGNDNICQDLISIT 784  
Db 725 DSVTPTILRLNFTLVGKPLAFLNLRPMLAALQRYFTASLPPFKNCGADHICQDNLGIS 784  
QY 785 RFSMSLDCLVVGPREENVTVVRNDGDSYRTQVTFPFLDLSYRKYSTLQNBQSRWS 844  
Db 785 RFPPLGKSLVGSNLELNAEVMVNDGDSYRTTTFSPAGLSYRYVAEGKQOQLRSL 844  
QY 845 RIACSSASTVSGALKSTCSINHPITPENSEVTNITFDVDSKASLGNKLLKANVTS 904  
Db 845 HLTCDSAPVG--SQGTWSTSCRIINELIFRGGAQITFLATEDVSPKAVLGRLLLTANVSS 902  
QY 905 ENNMERTKTEPOLRPLKVAIVMVTWTSHTVSTKYLNTAS--ENTSRVMOHQYQVSNLQ 963  
Db 903 ENNTPTSKTTFQLELPKVAIVTVVSSHQFTKYNFSESEKESHVAMERYQVNNIGQ 962  
QY 964 RSLPISLVFLVPLNQTIVWDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKPVPVC 1023  
Db 963 RDLFVSNFVWVPLNQEAVMVDVESHVPQNPFLSCSEKIAAPPASDFLAHIQKNVLDLC 1022  
QY 1024 STAVQRTQCDIPPGIOEEFNATLKNLSFDWYIKTSHNHLIVSTRAILFNDSVFTLL 1083  
Db 1023 STAGCLRFRCDFVPSVQSELDFTLLGNLSFGWVRQILQKKVSVVAEITFDTSVSQL 1082  
QY 1084 PQQAFVRSQTEKVEPEVFNPLPLVGVSSVGGLLLALITAAVLYKLGFFKROVDMWS 1143  
Db 1083 PQQAFVRSQTEKVEPEVFNPLPLVGVSSVGGLLLALITAVLYKLGFFKROVDMWS 1142

QY 1144 E 1144  
Db 1143 E 1143

## RESULT 4

ITAD HUMAN STANDARD; PRT; 1162 AA.  
AC Q13349; Q15575; Q15576;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (AD82).  
GN ITGAD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=96111956; PubMed=8777714;  
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,  
Stanton D.E., Gallatin W.M.;  
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3";  
RL J. Biol. Chem. 275:8959-8969(2000).  
RN [2]  
RP SEQUENCE OF 1-235 FROM N.A.  
RX MEDLINE=20187620; PubMed=10722744;  
RA Noti J.D., Johnson A.K., Dillon J.D.;  
RT "Structural and functional characterization of the leukocyte integrin gene CD11d. Essential role of Sp1 and Sp3";  
RL J. Biol. Chem. 275:8959-8969(2000).  
RN [3]  
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.  
RX MEDLINE=96257236; PubMed=8666289;  
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;  
RT "Cloning and chromosomal localization of a novel gene encoding a human beta 2-integrin alpha subunit";  
RL Gene 171:291-294(1996).  
RN [4]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99059842; PubMed=9841932;  
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W., Hoffman P.A., Stanton D.E., Bochner B.S.;  
RT "alpha beta 2 integrin is expressed on human eosinophils and functions as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-1)";  
RL J. Exp. Med. 188:2187-2191(1998).  
RN [5]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99370002; PubMed=10438935;  
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A., Grayson M.H., Bochner B.S., Gallatin W.M., Stanton D.E.;  
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a binding interface between I domain and VCAM-1";  
RL J. Immunol. 163:11984-11990(1999).  
CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-BORNE PATHOGENS, PARTICULATE MATTER, AND SENESCENT ERYTHROCYTES FROM THE BLOOD.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES, FOAM CELLS WITHIN ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.



CC	-/- SIMILARITY: Contains 1 VWFA domain.	QY	1	MALRVLLTALTTLCHGNFLDTENAMTQENRARGQSVVQLQSGSRVVGADQEIIVANQR	60
CC	-/- SIMILARITY: Contains 7 FG-GAP repeats.	DB	1	MTLXALLVLTALACHGNFLDTHTHPMTFQENAKGQGVVQLGGTVVVAAPQKAVNQI	60
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	QY	61	GSLVQCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVK	120
CC		DB	61	GALVQCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVK	120
CC		QY	121	GLCFPLGSLNLRQOQKPPBALRGCPQEDSDIAFLIDSGSIIIPHDPRMKFVSTVWEQL	180
CC		DB	121	GLCYLFGSLNLRPPQPPALRECEQESDITVFLIDSGSINNIDFQOKKFFVSTVWEQF	180
CC		QY	181	KSKTLPFLMOYSBEPRHFTFKFQNNPNRSLVKPITQLLGRTHATGRKVVRELFN	240
CC		DB	181	KSKTLPFLMOYSBEPRHFTFKFQNNPNRSLVKPITQLLGRTHATGRKVVRELFN	240
CC		QY	241	ITNGARKNAFLVITDGEKFGDPLGYEDVIPADREGVIRYVIGVGDARFSEKSRGL	300
CC		DB	241	ITNGARENAAKILVITDGEKFGDPLGYEDVIPADREGVIRYVIGVGDARFSEKSRGL	300
CC		QY	301	NTIASKPPRDHVFQNNFEALKTIQNLREKI FAIEGTQTCSSSSFEHEMSQEGPSAAIT	360
CC		DB	301	DTIASKPAGEHVFOVDNFEALNTIQNLQOEKI FAIEGTQTCSTSSFEHEMSQEGFSASIT	360
CC		QY	361	SNGPLLSVGSYDWMAGVFLYTSKEKSTFFINMTRVDSMDNDAYLGYAAAIILNRVQSLV	420
CC		DB	361	SNGPLLSVGSYDWMAGVFLYTSKEKSTFFINMTRVDSMDNDAYLGYAAAIILNRVQSLV	420
CC		QY	421	LGAPRYOHIGLVAMFRONTGMWESNANVKGTCIGAYFGASICSVDVDSNGSTDLVLICAP	480
CC		DB	421	LGAPRYOHIGLVAMFRONTGMWESNANVKGTCIGAYFGASICSVDVDSNGSTDLVLICAP	480
CC		QY	481	HYEYQTRGQGVSVCPPLPRGQARWQCDVAVLGEQGPWGRFGAALTVLGDVNGDKLTDVA	540
CC		DB	481	HYEYQTRGQGVSVCPPLPRGQARWQCDVAVLGEQGPWGRFGAALTVLGDVNGDKLTDVA	540
CC		QY	541	ICAPGEDNRGAVYLFHGTSGSGISPSHSQRIASKUSLPRLOYFGQSISGQDLTMDGLV	600
CC		DB	540	ICAPGEQENQGVYIFYGASIASASASHSHRIIGAHFSPGLQYFGQSISGQDLTMDGLM	599
CC		QY	601	DLTVGAQGHVLLRSQPVLRVKALMEFNPVARNVPCNDQVVKGEAGEVRVCLHVQK	660
CC		DB	600	DLAVGAQGHVLLRSQPVLRVKALMEFNPVARNVPCNDQVVKGEAGEVRVCLHVQK	659
CC		QY	661	STRDLREGQISVYVYDIALDSCRPHSRAVNETQNSTRTQTVLGLTQTCETILKLOLP	720
CC		DB	660	NTKDLREGQISVYVYDIALDSCRPHSRAVNETQNSTRTQTVLGLTQTCETILKLOLP	719
CC		QY	721	NCIEDPVSPVILRLNFSIAGTPLSAFGLNLRPVLAEDAORLFTALPPPEKNCQNDNICODD	780
CC		DB	720	DCVDSVSPVILRLNFSIAGTPLSAFGLNLRPVLAEDAORLFTALPPPEKNCQNDNICODD	779
CC		QY	781	LSITFSMSLDCLVVGGRPZFNVTVTVDNCGEDSYRTQVTFPFLDLISYRKVSTLQORS	840
CC		DB	780	LSITFSAMGLDTLVVGGRPQDNMSVTLRNDGSDSYRTQVTFPFLDLISYRKVSTLQORS	839
CC		QY	841	QBSWRL-ACESASSTENSGALKKSTCSINHPIDPENSEVTNITFDVDSKASLGNKLLK	899
CC		DB	840	KKPMFVKPAESSSSSEGGHGLKSTTNINHPIDPENSEVTNITFDVDSHASFNGKLLK	899
CC		QY	900	ANVTSENMPRTNTEFQLELPVKYAVVWVVTSHGVSTKYLNFTASENVTSMVQHYOVS	959
CC		DB	900	ATVASENNMSRTHKTKQLELPVKYAVVWVVTSHGVSTKYLNFTASENVTSMVQHYOVS	959
CC		QY	960	MLGQSLPISLVPLVPLVQLQTVIWDROPVTFSENLASTCHTKELPSSHFLAELRAP	1019
CC		DB	960	MLGQSLPISLVPLVPLVQLQTVIWDROPVTFSENLASTCHTKELPSSHFLAELRAP	1019
CC		QY	1020	VVNCISAVCQRIQCDIPFFGIGQEEFNATLKNLSFDWYIKTSHNELIVSTAEILFNDSV	1079
CC		DB	1020	VVNCISAVCQRIQCDIPFFGIGQEEFNATLKNLSFDWYIKTSHNELIVSTAEILFNDSV	1079

Query Match 76.3%; Score 4545; DB 1; Length 1153;  
Best Local Similarity 74.5%; Pred. No. 9.5e-295;  
Matches 860; Conservative 141; Mismatches 151; Indels 2; Gaps 2;

Query Match		99.7%; Score 5940.5; DB 1; Length 1152;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 1; Gaps 1;		
QY	1	MALRVLTLTALTLCHGNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNR 60
DB	1	MALRVLTLTALTLCHGNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNR 60
QY	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTRYK 120
DB	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTRYK 120
QY	121	GLCFPLGSLNLRQOPKPFPPALRGCPQEDSDIAFLIDSGSIIPHPFRMKRPVSTVMEQL 180
DB	121	GLCFPLGSLNLRQOPKPFPPALRGCPQEDSDIAFLIDSGSIIPHPFRMKRPVSTVMEQL 180
QY	181	KSKTLFSLMOYSSEBRFHFTFKFQONNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB	181	KSKTLFSLMOYSSEBRFHFTFKFQONNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
QY	241	ITNGARKNAFKLLVITDGEKFGDPLGYEDVTPADREGVIRYVIGGDAPRSEKSRQEL 300
DB	241	ITNGARKNAFKLLVITDGEKFGDPLGYEDVTPADREGVIRYVIGGDAPRSEKSRQEL 300
QY	301	NTIASKPPDRHVFQVNNFEALKTIONLREKIPALEGTQTGSSSSFEHMSQEGFSAIT 360
DB	301	NTIASKPPDRHVFQVNNFEALKTIONLREKIPALEGTQTGSSSSFEHMSQEGFSAIT 360
QY	361	SNGLPLSTGVSDNAGGVFLTSKSTFINTRVDSMDNAYLYGAAAILLRNVQSLV 420
DB	361	SNGLPLSTGVSDNAGGVFLTSKSTFINTRVDSMDNAYLYGAAAILLRNVQSLV 420
QY	421	LGAPRYQHIGLVAMFRONTKWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAP 480
DB	421	LGAPRYQHIGLVAMFRONTKWNESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLICAP 480
QY	481	HYETRTGGQSVCPPLRGQARWOCDAVLIGEQQQWGRFGAALTVLGDVNGDKLTJVA 540
DB	481	HYETRTGGQSVCPPLRGQARWOCDAVLIGEQQQWGRFGAALTVLGDVNGDKLTJVA 539
QY	541	IGAPEENNRGAVLYFHTGSSGSPSHSRIAGSKLSPRLQYFGQSGGQDITMDGLV 600
DB	540	IGAPEENNRGAVLYFHTGSSGSPSHSRIAGSKLSPRLQYFGQSGGQDITMDGLV 599
QY	601	DLTVGAQGHVLLRSQPLRVKAIMFNPVRVARNVPCNDQVVKGEAGEVRVCLHVQK 660
DB	600	DLTVGAQGHVLLRSQPLRVKAIMFNPVRVARNVPCNDQVVKGEAGEVRVCLHVQK 659
QY	661	STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQTCETLKLQLP 720
DB	660	STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQTCETLKLQLP 719
QY	721	NCIEDPVSPIVLRNFSVLGTPLSAFGNLRPLVLAEDAQRFLTALFPFKKNCNDNI QDD 780
DB	720	NCIEDPVSPIVLRNFSVLGTPLSAFGNLRPLVLAEDAQRFLTALFPFKKNCNDNI QDD 779
QY	781	LSITFSFSLDCLVGGPRFNVTVTVNDGEDSDVRTQVTFPFLDLSYKRVSTLQORS 840
DB	780	LSITFSFSLDCLVGGPRFNVTVTVNDGEDSDVRTQVTFPFLDLSYKRVSTLQORS 839
QY	841	QRWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900
DB	840	QRWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 899
QY	901	NVTSENMPRNKTEFQLELPVKVAVVMTSHGVSTKYLNFNTASENSTRVMOHQYQVSN 960
DB	900	NVTSENMPRNKTEFQLELPVKVAVVMTSHGVSTKYLNFNTASENSTRVMOHQYQVSN 959
QY	961	LGQSLSLPISLPLVPLVRLNQTVIWDROQVTESENLSSTCHTKERLPHSDFLAELRKPV 1020
DB	960	LGQSLSLPISLPLVPLVRLNQTVIWDROQVTESENLSSTCHTKERLPHSDFLAELRKPV 1019
Query Match		99.7%; Score 5940.5; DB 1; Length 1152;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 1; Gaps 1;		
QY	1021	VNCSIAVCQRTQCDIPFGIOEBFNATLKGNSLSPDWIKTSHNHLIIVSTABILENDSVF 1080
DB	1020	VNCSIAVCQRTQCDIPFGIOEBFNATLKGNSLSPDWIKTSHNHLIIVSTABILENDSVF 1079
QY	1081	TLLPQOGAFVRSQETKVEPEVENPFLIVGSSVGGLLALITAAALYKLGFFKQYKD 1140
DB	1080	TLLPQOGAFVRSQETKVEPEVENPFLIVGSSVGGLLALITAAALYKLGFFKQYKD 1139
QY	1141	MMSEGGPPGAPQ 1153
DB	1140	MMSEGGPPGAPQ 1152
RESULT 2		
ID	ITAM_MOUSE	STANDARD; PRT; 1153 AA.
AC	POS555;	
DT	01-NOV-1988 (Rel. 09, Created)	
DT	01-FEB-1991 (Rel. 17, Last sequence update)	
DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI).	
GN	ITGAM.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	SEQUENCE FROM N.A.
RP	MEDLINE=88312584; PubMed=3044779;	
RC	Pytela R.;	
RA	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";	
RT	EMBO J. 7:1371-1378(1988).	
EL	[2]	SEQUENCE OF 11-45 FROM N.A.
RN	STRAIN=BALB/c; TISSUE=Spleen;	
RP	MEDLINE=86287312; PubMed=2942940;	
RC	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.B.,	
RA	Larson R.S., Roberts T.M., Springer T.A.;	
RT	"A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).	
RN	[3]	SEQUENCE OF 17-28.
RP	MEDLINE=85198276; PubMed=3887182;	
RC	Springer T.A., Teplow D.B., Dreyer W.J.;	
RA	"Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";	
RT	Nature 314:540-542(1985).	
RL	Nature 314:540-542(1985).	
CC	-1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3. THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRANULATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.	
CC	-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.	
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.	
CC	-1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.	
CC	-1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.	
CC	-1- SIMILARITY: Belongs to the integrin alpha chain family.	

RP SEQUENCE OF 1-9 FROM N.A.  
RX MEDLINE=92073318; PubMed=16683702;  
RA Shelley C.S., Arnaut M.A.;  
RT "The promoter of the CD11b gene directs myeloid-specific and  
RL developmentally regulated expression.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).  
[7]  
RP SEQUENCE OF 1-9 FROM N.A.  
RC TISSUE=Blood;  
RX MEDLINE=92144986; PubMed=1346576;  
RA Paul H.L., Rosmarin A.G., Tenen D.G.;  
RT "Characterization of the myeloid-specific CD11b promoter.";  
RL Blood 79:865-870(1992).  
[8]  
RP SEQUENCE OF 17-31.  
RX MEDLINE=87076671; PubMed=3539202;  
RA Pierce M.W., Remold-O'Donnell E., Todd R.P. III, Arnaut M.A.;  
RT "N-terminal sequence of human leukocyte glycoprotein Mol.  
RL conservation across species and homology to platelet IIB/IIIA.";  
RL Biochim. Biophys. Acta 874:368-371(1986).  
[9]  
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.  
RX MEDLINE=95171458; PubMed=7867070;  
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;  
RT "Crystal structure of the A domain from the alpha subunit of integrin  
RL CR3 (CD11b/CD18).";  
RL Cell 80:631-638(1995).  
[10]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.  
RX MEDLINE=96363671; PubMed=8747460;  
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;  
RT "Two conformations of the integrin A-domain (I-domain): a pathway for  
RL activation?";  
RL Structure 3:1333-1340(1995).  
[11]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.  
RX MEDLINE=98362595; PubMed=9687375;  
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,  
RA Fairbanks M.B., Pinzel B.C., Garlick R.L., Heinrikson R.L.,  
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,  
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;  
RT "Cation binding to the integrin CD11b I domain and activation model  
RL assessment.";  
RL Structure 6:923-935(1998).  
[12]  
RP 3D-STRUCTURE MODELING OF 17-616.  
RX MEDLINE=98226734; PubMed=9560195;  
RA Oxvig C., Springer T.A.;  
RT "Experimental support for a beta-propeller domain in integrin alpha-  
RL subunits and a calcium binding site on its lower surface.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).  
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS  
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES  
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.  
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF  
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D  
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR  
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES  
CC OF FIBRINOGEN GAMMA CHAIN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M  
CC ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND  
CC GRANULOCYTES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VMPA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- DATABASE: NAME-PROW, NOTE-CD guide CD11b entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; J03925; AAA59544.1; -  
DR EMBL; M18044; AAA59491.1; -  
DR EMBL; J04145; AAA59903.1; -  
DR EMBL; S5227; AAB34821.1; -  
DR EMBL; S52152; AAB34821.1; JOINED.  
DR EMBL; S52153; AAB34821.1; JOINED.  
DR EMBL; S52154; AAB34821.1; JOINED.  
DR EMBL; S52155; AAB34821.1; JOINED.  
DR EMBL; S52157; AAB34821.1; JOINED.  
DR EMBL; S52159; AAB34821.1; JOINED.  
DR EMBL; S52161; AAB34821.1; JOINED.  
DR EMBL; S52164; AAB34821.1; JOINED.  
DR EMBL; S52165; AAB34821.1; JOINED.  
DR EMBL; S52167; AAB34821.1; JOINED.  
DR EMBL; S52169; AAB34821.1; JOINED.  
DR EMBL; S52170; AAB34821.1; JOINED.  
DR EMBL; S52173; AAB34821.1; JOINED.  
DR EMBL; S52174; AAB34821.1; JOINED.  
DR EMBL; S52180; AAB34821.1; JOINED.  
DR EMBL; S52181; AAB34821.1; JOINED.  
DR EMBL; S52184; AAB34821.1; JOINED.  
DR EMBL; S52185; AAB34821.1; JOINED.  
DR EMBL; S52191; AAB34821.1; JOINED.  
DR EMBL; S52192; AAB34821.1; JOINED.  
DR EMBL; S52203; AAB34821.1; JOINED.  
DR EMBL; S52212; AAB34821.1; JOINED.  
DR EMBL; S52213; AAB34821.1; JOINED.  
DR EMBL; S52215; AAB34821.1; JOINED.  
DR EMBL; S52219; AAB34821.1; JOINED.  
DR EMBL; S52220; AAB34821.1; JOINED.  
DR EMBL; S52221; AAB34821.1; JOINED.  
DR EMBL; S52222; AAB34821.1; JOINED.  
DR EMBL; S52226; AAB34821.1; JOINED.  
DR EMBL; M76724; AAA58410.1; -  
DR EMBL; M84477; AAA51960.1; -  
DR PIR; A31108; RHUUB.  
DR PDB; 1ABX; 17-JUN-98.  
DR PDB; 1BHQ; 18-NOV-98.  
DR PDB; 1BHQ; 18-NOV-98.  
DR PDB; 1IDN; 25-NOV-98.  
DR PDB; 1IDO; 01-AUG-96.  
DR PDB; 1JLM; 11-JAN-97.  
DR PDB; 1MIU; 07-AUG-02.  
DR Genew; HGNC:6149; ITGAM.  
DR MIM; 120980; -  
DR GO; GO:0008305; C:integrin complex; TAS.  
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
DR GO; GO:0007155; P:cell adhesion; TAS.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 4.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; 3D-structure; Repeat; Magnesium; Calcium.  
FT SIGNAL 1 16  
FT CHAIN 17 1152 INTEGRIN ALPHA-M  
FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL)  
FT TRANSMEM 1105 1128 POTENTIAL  
FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 10.1123 Seconds  
(without alignments)  
5937.039 Million cell updates/sec

Title: US-09-902-481b-1  
Perfect score: 5956  
Sequence: 1 MALRVLLTALTLCGFLND.....FKROYKDMSEGPGCAEPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5940.5	99.7	1152	1 ITAM_HUMAN	P11215 homo sapien
2	4545	76.3	1153	1 ITAM_MOUSE	P05555 mus musculus
3	3503	58.8	1163	1 ITAD_HUMAN	P20702 homo sapien
4	3455	58.0	1162	1 ITAD_HUMAN	Q13349 homo sapien
5	1563	26.2	1170	1 ITAL_MOUSE	P20701 homo sapien
6	1549	26.0	1163	1 ITAL_MOUSE	P24063 mus musculus
7	1165.5	19.6	1167	1 ITAE_MOUSE	Q60677 mus musculus
8	1163	19.5	1179	1 ITAE_HUMAN	P38570 homo sapien
9	1108.5	18.6	1151	1 ITAH_HUMAN	P56199 homo sapien
10	1102.5	18.5	1189	1 ITAH_BOVIN	O9UKX5 homo sapien
11	1099	18.3	1170	1 ITA2_MOUSE	P53710 bos taurus
12	1089	18.3	1178	1 ITA2_MOUSE	Q62469 mus musculus
13	1087.5	18.3	1181	1 ITA2_HUMAN	P17301 homo sapien
14	1085.5	18.2	1167	1 ITAG_HUMAN	O75578 homo sapien
15	1082	18.2	1180	1 ITAI_RAT	P18614 rattus norv
16	667	11.2	1039	1 ITA4_MOUSE	Q00651 mus musculus
17	638	10.7	1038	1 ITA4_HUMAN	P13612 homo sapien
18	630	10.6	1035	1 ITA9_HUMAN	Q13797 homo sapien
19	600	10.1	1032	1 ITA4_XENLA	Q91687 xenopus lae
20	571.5	9.6	1066	1 ITA3_CRISP	P17852 cricetidae
21	567.5	9.5	1053	1 ITA3_MOUSE	Q62470 mus musculus
22	555.5	9.3	1053	1 ITA5_MOUSE	P11688 mus musculus
23	550	9.2	1034	1 ITAV_CHICK	P26008 gallus gall
24	546.5	9.2	1050	1 ITA5_XENLA	Q06274 xenopus lae
25	540	9.1	1130	1 ITA6_HUMAN	P23229 homo sapien
26	535	9.0	1044	1 ITAV_MOUSE	P43406 mus musculus
27	532	8.9	1049	1 ITA5_HUMAN	P08648 homo sapien
28	531.5	8.9	1066	1 ITA3_HUMAN	P26006 homo sapien
29	530.5	8.9	1072	1 ITA6_CHICK	P26007 gallus gall
30	526	8.8	1048	1 ITAV_HUMAN	P06756 homo sapien
31	517	8.7	1044	1 ITA8_CHICK	P26009 gallus gall
32	512.5	8.6	1091	1 ITA6_MOUSE	Q61739 mus musculus
33	508	8.5	1396	1 ITA2_DROME	P12080 drosophila

RESULT 1				
ITAM_HUMAN	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89089893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,			
RT	Roch G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	498	8.4	1179	1	ITA7_MOUSE	Q61738 mus musculus
35	492	8.3	1033	1	ITAB_MOUSE	Q9QUM0 mus musculus
36	491.5	8.3	1146	1	ITAI_DROME	Q24247 drosophila
37	489.5	8.2	1039	1	ITAB_HUMAN	P08514 homo sapien
38	489	8.2	1025	1	ITAB_HUMAN	P53708 homo sapien
39	486	8.2	126	1	ITAM_CAVPO	P11578 cavia porce
40	474.5	8.0	1226	1	PAT2_CAEEL	P34446 caenorhabdi
41	472	7.9	1181	1	ITAT_HUMAN	Q13683 homo sapien
42	469.5	7.9	1106	1	ITAT_RAT	Q63258 rattus norv
43	454.5	7.6	1139	1	INAI_CAEEL	Q03600 caenorhabdi
44	424.5	7.1	1115	1	ITAI_DROME	O44386 drosophila
45	383	6.4	1000	1	ITAS_DROME	Q9WIM8 drosophila

ALIGNMENTS

Db 702 DFL-----KCSVGFPPMRSSKYEFSVDPDTSHLS--GEEVLSPIVTAQSG--NTER 750  
 Qy 914 TE-----FQLELPVKYAVVWVTS-----HGVSTKYLNFTASNTS---RVMQHOY 956  
 Db 751 SESLHDVTLVLMVMEHVDTSITGIMSPTSFVIGESVDAAFIQLDLCHFPQINIL 810  
 Qy 957 QVSNLQSRSLPISLVL-VPVRLN-----QTVWDRPQVPTFSENLSSTCHTKER 1004  
 Db 811 QVYATGPSTLPGSSVSISFPNLSGGGAEMPHVQEMVVGQKGNCSFQKNPTPCIIPOEQ 870  
 Qy 1005 LPSHSDFLAELRKA-----PWNCSIAVCQRIQCDIPFPFGIQBEFNATLKNLSPD 1055  
 Db 871 ENIFHTTFAPTKGRKVLDCKEFGISCLTACHN-----PSLAKEBSEKTI-----D 917  
 Qy 1056 WYIKTSHNHLIVSTAEILFNDSTVFTLLPGQCAFVRSQTERKVP----- 1100  
 Db 918 IY-----MLLAT-BELKKDSSVIQ-----FMSRAKVKVDPALRVVEIAHGNPEEV 962  
 Qy 1101 ---PEVPN---PLPLIVG-----SSVGLLLALITAAIKLGPFRQYKMM 1142  
 Db 963 TVVFEALHNLPRGVVGVWIIAISLLVIGILIFLLAVLLWMQGFRRRYKEII 1015

RESULT 15  
 T31437  
 integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)  
 C:Species: Lytechinus variegatus (variegated urchin)  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T31437  
 R:Hertaler, P.L.; McClay, D.R.  
 submitted to the EMBL Data Library, May 1998  
 A:Description: Alpha SU2, a sea urchin integrin which binds laminin.  
 A:Reference number: 221035  
 A:Accession: T31437  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1041 <HER>  
 A:Cross-references: EMBL:AF067658; NID:G3220240; PID:G3220241; PIDN:AAC23572.1  
 A:Experimental source: developmental stage embryo  
 C:Function:  
 A:Description: binds laminin  
 C:Superfamily: integrin alpha-2b chain

Query Match 10.3%; Score 614.5; DB 2; Length 1041;  
 Best Local Similarity 24.8%; Pred. No. 4e-34;  
 Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

Qy 352 QEGFSALITNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDS-----DNNDAYLG 405  
 Db 181 QAGSGIIFSDNSALVNGAFSGSYLQGIYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239  
 Qy 406 YAAA1--ILNRVQSLVLAGAPRYOHI-GLVAMPFRONTGMWESNANVKGTQIGAYFGASLC 462  
 Db 240 YSLALGDFNGDGVQDYVGTFRABSLMGLVAIFQNLNQFN---QVMSTQIVAYFGYSVT 296  
 Qy 463 SVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCPPLRQORARWQCDVLYGEQ----- 514  
 Db 297 VVDI--NNDTYDILLVAGAPMYMDGPAIQ-----RWEAGAVVYVLQNPVGPGA 343  
 Qy 515 -----GQWGRFGAALTVLGDVNGDKLTVAIGARGEEDNRGAVLPHGTSGS 562  
 Db 344 SNLSLSSTLIGQIRERFGLSTASIGDSNQDGFNDVAIGAPYEGDAGAVIYHG--SAN 402  
 Qy 563 GISPSHSQRIAGSKLS-PRIQYFGQSLSGGQDLTMDGLVLTVGAQ--GHVLLRSQPV 619  
 Db 403 GLKSTPAQVLTPTSLGHSGITTFGFSLQGGQDMKKNYPDLLVGAESANTAVLIRTPVV 462  
 Qy 620 RVKALMEFNRVAVRNVPCNDQVWKEAGEVRVCLHVQKSTEDRLREGIQSV----- 674  
 Db 463 SLDATLNTPE-----IGINLENKTYB-LADGTMTVTSFIAMT 497  
 Qy 675 -----VTYDLDLDSG-RPHSRAVFNETKNSRROTQVIGL--TOTCEILKLQ 718

Search completed: June 7, 2004, 17:17:33  
 Job time : 20.7639 secs

Db 498 CFTYTGNYLPDHIDISVTVTVDSGIIANRRAMFVDNDMSEITKTRRLAVSTQPCDPLRAY 557  
 Qy 719 LPNCIEDPVSPIVLRNFSLVCTPLSAGFN-----LRPVLAEDAQRLFTALPFEK 769  
 Db 558 VGNISIEDKUTPTKVITYQIDL-----NNDSERLQPHLPIIDMATSTQTKQVSIQN 609  
 Qy 770 NCGNDICODDLISITFSFMSLDCLVVGGPREFNVTVVRNDGSDSYRTQVTFPPDLDSY 829  
 Db 610 MCVN-NICIPDLVDVT-PNLPNIVIGQTOELTLDVSLANRGEDAFQSSLSVYVPLGLQF 667  
 Qy 830 RKVSTLQNRQSRWKLACESASSTEVSGALKSTCSINHPIPPEN-----SEVFNIT 893  
 Db 668 VRL-----ERKANMDFVTCSESD-----LRITCDTGNPMVGRKILBFGLTLSTFQVS 717  
 Qy 884 PDVDS-----KASLGNKLLKLLKANVTSENMMPRINKTEFOLEFPVKYAVVMTSHGVSTK 938  
 Db 718 GDKDSIEPTFKABSENS--EDPNTLENNELNMTVPVTVDTCLKLSASYPEIWMYSTQED 775  
 Qy 939 YL-----NFTASENTSRVMOHQYQV-----SNLQORSILPIS-----LVFLVPV 976  
 Db 776 YVVPFPKAKNASEADIGMEVM-HLYEVRNMGSGNAGEVSLNIQWPOKNEDEGEYLFYLLGI 834  
 Qy 977 RLNQTVIWDPRQ-----VTFSENLSSTCHTKERLPSESDPLAELRKAPVV 1021  
 Db 835 MTEEGVTCOLTOGKANPEGVKLEPSTKAKLSNSTTQVSGRKRPEPEVAEALQTDN--VI 892  
 Qy 1022 NCSIAVCQRIQCDIPFPFGIQBEFNAT-----LKGNLSTFDWYIKTSHNHLIVSTAEILF 1075  
 Db 893 YCASDSCVLINCTI-----DEINASKSVVAILGRF--W-----ERTF 928  
 Qy 1076 NDSVFTLLPGQCAFVRSQTERKVE--PPEVPNP-----LP----- 1108  
 Db 929 QKAVSELTPWQATIASASAAVKTIPIPLPRDPSDSTKASTIVTTEELVPPVTPIAW 988  
 Qy 1109 -LIVGSGVGLLLALITAAIKLGPFR-----QYKDMMS--EGGPP 1148  
 Db 989 WIIVSVLGGIILILLIILGLWKCGFFERKPKGEKEVAPVASADKGGP 1038









Qy	178	EOLX--KSKYLPSLMOYSEBPRHETFEKQONNPNRSLVKPITOLL-----GRTHATGI	231
Db	189	QGLDIGETKTMGLIQIYANNPRVFNLTNPKSD--EMIKATSQTFQYQGDULTNFKAI	245
Qy	232	RKVRELFNTNGARKNAFKLLVITDGEKFGDPLGYEDVTPBADRGVIRYVIGV----	287
Db	246	QYARDTAYSTAAAGGPGCATKVVVVVTDGESH--DGSKLKAVIDOCNKONILPGVLAVLYL	304
Qy	288	GDAPFSEKRSQBLNTTASQPRDHVTFQVNNFBAKTIQNLREKFIPIAIGTOTGSSSSP	346
Db	305	NRNALDFTNLIKETIKATASIPTEPHFPFNVSDBADLLEKAGTIGEQIFSIETGVQG--	363
Qy	347	EHEMSQBGFAAIT--SNGPLLSVTGSDWAGGVFLYTSKEKSPFINMT--RVDSDKN-D	401
Db	364	QMEKSQVGFABAYSPPONILMLGAVGAYDWSGVVQKTPHGHLIFSQAQEUILQDRNHS	423
Qy	402	AYLGAYAAAILNRVQSLVIGAPRYQHILGVAMPQNTGMWESNAVY-----KGTGIGA	455
Db	424	SYLGVSVASISTGNSVHFVAGAPRANTGTQIIVSYVN-----ENGVTVTQSQRGDQIGS	478
Qy	456	YFGASICSVDVDSNGSTDVLIGAPHYYEQTR--GGQVSVCPILPRQARAWQCDVAVLGE	513
Db	479	YFGSVLCAVDVNXDTITDVLVGAPMTMDLKEEGRVYLFTITKG--ILNKH--QFLEGP	535
Qy	514	QCGPWGFGAALTIVLGVNDGKLTDAIGAPBEDNRGAVYLFHGTSGSISPSHSORIA	573
Db	536	NGLENARFGSAIAALSIDINMDGFNDVTFVGSPLENQNSGAVIYNGHEGM--IRLRSOKIL	594
Qy	574	GS--KLSPLRLQYFCQSLSGGQDLTMDGLVBLTVGAGQGHVILLRSQPVLYRKAIWEFNPRE	631
Db	595	GSDRAPFSSHLOYFGRSLDVGDLNGDSITDVSUGAFQVVQLMWSQSADVSVVDASFTPKK	654
Qy	632	VARNVFCNDQVVKGEAGEVRLVHVQKSTRDRMLREGQIQSVVYTDLALD----SGRPH	687
Db	655	I--TLINKNAEI-----KLKCF--SAKFRPTQNNQVATVYITITDEQFSSRVI	701
Qy	688	SRAVFNETKNSRTOCVLGLTQICE--TLKQLQPNCEIDPVSPVLRLNFSI--VCTPL	743
Db	702	SRGLFKENNERCLOKTMIVSQACRSYIIRHQEPS--DIISPLNOMISLENPGT--	756
Qy	744	SAPGNLRLVLAEDAQRLFTALFPFERKCGNDNICQDDLSTTF-----SFMSLDCLVVGGR	799
Db	757	-----NPALFAYSETVKVFSIPPHKCGDDBGVCSIDLVLNVQQLPATQOQPFIVSNQNK	810
Qy	800	EFNVTVTRVNDGDSYRTQVTFPPFPDLVYRKVSTLQNRQSQRWLACESAST--RVSG	858
Db	811	RLTFSVQLKNGKESAYNTEIIVDSENLFF-----ASWSMPVDTGTEVTCQIAS	858
Qy	859	ALKSTCSINHPIDPENSEVTFTNITPVDKASIGNKLLKANVTSENMPRTNKTBFOL	918
Db	859	SQKSVTCNVGVPALKSKQOVFTLINFDFNLQ--NLQNASISFRALSSQENMADNSVNL	917
Qy	919	ELPVKYAVMVVTSHGVSCTKLNFNTASENTRVMHQYQVSNLQOR-----SLPI	968
Db	918	KLSELYDAEIHIT--RSTNINPFVSLDGNVSSV--HSFE--DIGKPTFISKVITGSPVP	973
Qy	969	SLVELVPRLMQTVIWRDPQVTPSEN--LSSTCHTKE-----RLPSHS	1009
Db	974	SMA-----SVLIHPQVTKDKNPLMYLTGVHTDQAGDISCEASINPLKIQOTSSV	1024
Qy	1010	DPLAE-LRKAPVNCSTAVCORICQDIPFGIQIEFNATLKNLSFDWYIKTSENHLLIV	1068
Db	1025	SFXSENFRHITKELNCRATASCNINCMWLRLDQVKGSEYFLNSTRINWGTFAASTFQTVOLT	1084
Qy	1069	STABE-LPNDVSVFTL-----LPQCGAFVRSOTETKVPPEB-VBNPLVLIVSGSVGLLL	1120
Db	1085	AAAEIDYNQIYVIEBNTVIP-----LTIKPKHEKVEVPGVIVGSVIAGILL	1134
Qy	1121	LALITAILYKLGFPPKQYKOM	1141
Db	1135	LLALVAILWLKLGFFPKRYKYM	1155

Qy	178	EOLX--KSKYLPSLMOYSEBRHETHTPEKQONNPNRSLVKPITOLL-----GRTHATGI	231
Db	189	QGLDIGETKTMGLIQUANNPRVFNLTNPKSD--EMIKATSQTFQYQGDULTNFKAI	245
Qy	232	RKVRELFNTNGARKNAFKLLVITDGEKFGDPLGYEDVTPBADRGVIRYVIGV----	287
Db	246	QYARDTAYSTAAAGRGPGATKVVVVVTDGESH-DGSKLKAVIDOCNKONILRPGVLAVLYL	304
Qy	288	GDAPFREKSRQBLNTTASRPRDHVTVQVNNFBAKTIQNLREKFIPIAIGTOTGSSSSP	346
Db	305	NRNALDPKNIKEIKATIASIPTEPHFPFNVSDBADLLEKAGTIGEQIFSIETGVQG-GDNF	363
Qy	347	EHEMSQBGFAAIT--SNGPLLSVTGSDWAGGVFLYTSKEKSPFINMT--RVDSDKAN-D	401
Db	364	QMEKSQVGFABAYSPPONILMLGAVGAYDWSGVVQKTPHGHLIFSQAQEUILQDRNHS	423
Qy	402	AYLGYAAAIILNRVQSLVIGAPRYQHILGVMPRQNTGMWESNAVY-----KGTGIGA	455
Db	424	SYLGSVASISTGNSVHFAGAPRANTGTQIVLYSVN-----ENGNTVTVQSQRGDOIGS	478
Qy	456	YFGASICSVDVDSNGSTDVLIGAPHYYEQTR--GGQVSVCPILRQARARQCDVAIYGE	513
Db	479	YFGSVLCAVDVNXDTITDVLVGAPMTMDLKKEBGRVYLFTITKG-ILNKH--QFLEGP	535
Qy	514	QCGPWGFGAALTIVLGVNDGKLTDAIGAPBEDNRGAVYLFHGTSGSISPSHSORIA	573
Db	536	NGLENARFGSAIAALSIDINMDGFNDVTVGSPLENQNSGAVIYNGEHGM-IRLRYSKIL	594
Qy	574	GS--KLSPLRLQYFCQSLSGGQDLTMDGLVBLTVGAGQGHVILLRSQVPLRYKAIWEFNPRE	631
Db	595	GSDRAPFSHLOYFGRSLDYGDLNGDSITDVSUGAFQVVQLMWSQSADVSVVDASFTPKK	654
Qy	632	VARNVFCNDQVVKGEAGEVRLVHVKQSTRDRMLREGQIQSVVYTDLALD----SGRPH	687
Db	655	I--TLINKNAEI-----KLKCF--SAKFRPTQNNQVATVYNTITDEQFSSRVI	701
Qy	688	SRAVFNETKNSRROTQVLGLTQICE--TLKQLQPNCEIDPVSPVILRLNFSI--VCTPL	743
Db	702	SRGLFKENNERCLOKTMIVSQACRSYIIRHQEPS--DIISPLNOMISLENPGT--	756
Qy	744	SAPGNLRPLAEDAQRLFTALFPFERKCGNDNICQDDLSTTF-----SFMSLDCLVVGGR	799
Db	757	-----NPALFAYSETVKVFSIPHHKCGDDGVCISDLVLNVQQLPATQOQPFIVSNQNK	810
Qy	800	EFNVTVTVRNDGDSYRTQVTFPPFPDLRYRKVSTLQNRQSQRWLACESAST-EVSG	858
Db	811	RLTFSVLQKNKGEAYNTEIIVDSENLFF-----ASWSMPVDTGTEVTCQIAS	858
Qy	859	ALKSTCSINHPIDPENSEVTFTNITPVDKASIGNKLLKANVTSENMPRTNKTBFOL	918
Db	859	SQKSVTCNVGVPALKSKQOVFTLINFDFNLQ-NLQNASISFRALSSQENMADNSVNL	917
Qy	919	ELPKVAYVMVVTSHGVSTKYLNTFASENTSRVMOHQVQVSNLQOR-----SLPI	968
Db	918	KLSELVDAEIHIT-RSTNINPFYSLDGNVSSV-HSFE--DIGKPFIFIKVTGTGVPV	973
Qy	969	SLVELVPRLMQTVIWRDPQVTPSEN--LSSTCHTKE-----RLPSHS	1009
Db	974	SMA-----SVLIHPQVTKDKNPLMYLTGVHTDQAGDISCEASINPLKIQOTSSV	1024
Qy	1010	DPLAE-LRKAPVNCSTAVCORICQDIPFGIQIEFNATLKNLSPDWYIKTSENHLLIV	1068
Db	1025	SFXSENFRHikelNCRATASCNINCMWLRLDQVKGIEYFLNSTRINWGTFAASTFQTQVLT	1084
Qy	1069	STABE-LPNDVSFTL-----LPQCGAFVRSOTETKVPPEB-VBNPLVLIVSGSVGLLL	1120
Db	1085	AAAEIDYNQIYVIEBNTVIP-----LTIKPKHEKVEVPGVIVGSVIAGILL	1134
Qy	1121	LALITAILYKLGFPRQYKOM	1141
Db	1135	LLALVAILWLKLGFFPRKYERM	1155

QY 1107 LPLTVGSSVGGHLLALITAAALYKLGFFKQYKDMSE 1144  
 Db 1125 LPLIIGSSVGGHLLALITAAALYKLGFFKQYKDMSE 1162

RESULT 7  
 A45226  
 integrin alpha-1 chain - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Sep-2003  
 C:Accession: A45226  
 R:Brizewitz, R.; Epstein, M.R.; Marcantonio, E.E.  
 J. Biol. Chem. 268, 2989-2996, 1993  
 A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit.  
 A:Reference number: A45226; MUID:9315124; PMID:8428973  
 A:Accession: A45226  
 A>Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-1151 <BRI>  
 A:Experimental source: hepatoblastoma cell line HepG2  
 A>Note: sequence extracted from NCBI backbone (NCBIP:124326)  
 F:142-317/Domain: von Willebrand factor type A repeat homology <WVAL>

Query Match 18.6%; Score 1108.5; DB 2; Length 1151;  
 Best Local Similarity 28.0%; Pred. No. 2e-68;  
 Matches 346; Conservative 210; Mismatches 486; Indels 195; Gaps 44;

QY 17 ENLDTENAMTFOENARG-FQGSVVOL---QGSVVVVGAPQEIIVANORGLYQCDYSTGS 72  
 Db 1 FNVDDKSNMTFGSPVEDMPGTYTQVYENBEGKVLGSLPGVQKQRTGVDYKCPVGRGE 60

QY 73 CEP-IRLOVPEA-----VNMSLGLSLAATTPPOLLACGPTVHOTCSENIVYKGL 122  
 Db 61 SLPCVKLDLPWNTSIPNVTVEKNMTFGSL-VTNENGFLACGLPYAYRGHLYTGI 119

QY 123 CFLFGNLRQOPKPPALRGCPQEDSDIAFLDGGSGIIPHRFRMKPFVSTVMEQLK- 181  
 Db 120 CSDVSPTFQVNSIAP--VOECSTQ-LDIVIVLDGNSIYPMD--SVTAFNLNLRMDI 174

QY 182 -KSKTLFSLMOYSEPRHFTHFTKFEONNPNRSLVPIPTOLLGR-THATGIRKVVREL 239  
 Db 175 GPKQTQGVIGYGENVTHEFNLYKSYSTEVLVAKKIVQRGGRQWTALGDTARKEAF 234

QY 240 NITNGARKNAFLVITGKEGDPGLVEDVTPADREGRVIRYIVGVGDAPR-----SE 294  
 Db 235 TEARGARRGVKKVWVITDGESH-DNHRLLKKVIQDCEDENIQRFSAIILGYSNRGNLSTE 293

QY 295 KSRQELNTIASKEPRDHPVQVNNFEALKTIONOLREKI FALGCTGTGSSSSSEHEMSQEG 354  
 Db 294 KFVEEIKSIASEPTEKHPFNVSDALVTLVKTGLGRIFALEATADQSAASFEMMSQTG 353

QY 355 FSAATISNGPLSTVGSYDWAGGVFLYTSKE-----KSTP-INMTVDSMDMDAYLGYAA 408  
 Db 354 FSAHYSDQVWMLGAVGAYDWGTVVMQKASQIIIPRNTTFNVESTKKNPL-ASYLGITYV 412

QY 409 AILRENVOSL-VLGRPRVQHIGLVAMFRQNTGWESNANVKGTOIGATFGASLCSVDV 457  
 Db 413 NSATASSGDVLYTAGOPRYNHTGQVIYRMEDGNIKILQTLSEQIGSYFGSILATTDID 472

QY 468 SNGSTDVLVLGAPHY-----YEQTR-GGQVSVCPPLPRGQARWQCDVAVLY 511  
 Db 473 KDSNTDILLVAGPMYWGTEKEOGKYVYVALNQTRFEXQMSLEPIKQYCCSRQHSCTT 532

QY 512 GQQCQPWG-RFGMALTVLDVNGDKLTVAGAPGEEDNRGAVYLFHGTSGSGISPSHSQ 570  
 Db 533 ENKNEPCGARFSTAIAAVKDLNLDGNDIVIGAPLEDHGGAVIYHG-SGRTIRKEVAQ 591

QY 571 RIAGSKLSPRLQVFGQSLGSGQDLTMDGLVDLVGAQGHVLLRSLQPLRVKALMEENPR 630  
 Db 592 RIFSGDGTALFFGQSIHGMNDLNGDGLTDTVIGLGGALFWERDVAIVKVTWTFNFBP 651  
 QY 631 EVARNVFECDNQVYKGEAG--EVRVCLHVQ-KSTRDLRLBQIQSVVYDVLALDSGRPH 687

Db 652 KVNIOKKNCH---MEKETVCINATVCFVKLSKEDTITYEADLQ-----YRVTLSLRQI 704  
 QY 688 SRAVFNET-----KNSTRQTVLGLTQTCETKLQLPNCIEDPVSPVLRINPFLVGT 741  
 Db 705 SRSFSGTQERKQVQVNIIVKSEC-----TKHSFYMLDKHDFQDSVR---IITLDFNLT-D 755

QY 742 ELISAFGNLRPVLAEADQRLFTALPFPKPKGNDNICQDDLSITFTSPMSLDCLVVGSPRE- 800  
 Db 756 PENG-----PVLDDSLPNSVHEIYIPAKDCGKKEKICISDLSHVATTEKDLIIIVRSQNDK 810

QY 801 FNVTVVRNDGDSYRTQVTPFPPLDLSYRKVSTLQVRSQSRWRLACESASSTEVSGAL 860  
 Db 811 FNVSLTVKTKDGAAYNTRIVHSPNLPVSGIEAQKD-----SCSEN----- 853

QY 861 KSTSCSINHPIFPENSEVTFNITFDVDSKASIGN-KLLKANVTSENNMPTNKTFQLE 919  
 Db 854 HMITCKVGYFFLERGEMVTFKILFQNTSYLMENVTIYLSATSDSEEPETLSDNVNIS 913

QY 920 LPVKYAVVMVTVSHGVSTKYNFTASENTRVMOHQYQVSN-----LQGRS-----L 966  
 Db 914 IPKVEVGLQFTS-SASEYHISIAANETVPEVINSTEDIGNINIFYLIRKSGSPMPDEL 972

QY 967 PISLVP-----LVPVRLNQTVIMDRPQVTFSENLSSTCTHKE-----RLPS 1007  
 Db 973 KLSISGPNMTSNGYPVLYPTGLSS-----SENANCRPHIFEDPFSINSGRKMTT 1021

QY 1008 HSDFLAELRKAPVNCSTAVCORIQCDIPFFQIE-----EFNATLK 1049  
 Db 1022 STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILWKPTFIKSYFSSLNLTIR 1078

QY 1050 GNLSPDMWIKTSHNHLIVSTAEIILFNDISVFTLLPGQAFVRSQTBTKVPEPEVPNPLP 1109  
 Db 1079 GEL-----REENASLVLSNN-----QKRELAIQISKGLGRVPL 1114

QY 1110 ---IVGSSVGGHLLALITAAALYKLGFFKQYKDMSE 1144  
 Db 1115 MWILLSAFAGLLLLALLALWKLIGFFRPLKKNMEK 1151

RESULT 8  
 I45914  
 integrin alpha 2 subunit - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Sep-2003  
 C:Accession: I45914  
 R:Kamata, T.; Puzon, W.; Takada, Y.  
 J. Biol. Chem. 269, 9659-9663, 1994  
 A:Title: Identification of putative ligand binding sites within the I-domain of integrin alpha 2 subunit.  
 A:Reference number: A54402; MUID:94193647; PMID:7511592  
 A:Accession: I45914  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1170 <KAM>  
 A:Cross-references: GB:I25886; NID:G439695; PIDN:AAB59255.1; PID:G439696  
 F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.5%; Score 1099; DB 2; Length 1170;  
 Best Local Similarity 27.8%; Pred. No. 9.6e-68;  
 Matches 340; Conservative 215; Mismatches 498; Indels 168; Gaps 47;

QY 11 LTLCHGFNLDENAMTFO-ENARGFQGSVVOL---QGSVVVVGAPQEIIVANORGLYQCD 66  
 Db 13 LNCVAVNVGLPFAKIFSGPSSEQFGYAVQVQFIPNKGWLLVGSFPGPKNKGVDVYKC 72

QY 67 --DYSTGSCBPIRLQ-----VPEAVNMSLGLSLAATTPPOLLACGPTVHOTCSEN 117  
 Db 73 PVDLSTTTCERKLNQVTSNMTVMKTNMSLGLTLFNVGVTGGFLTCGLPLWAOQGSQY 132

QY 118 YVKGCLFLFGNLRQOPKPPALRGCPQEDSDIAFLDGGSGIIPHRFRMKPFVSTV 177  
 Db 133 YTTGVCSVSDPF-QLRTSFAPVQTCF-SFIDVWVWCDENSIYPMD--AVKNFLKFEV 188





F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>  
F:20-1107/Domain: extracellular #status predicted <EXT>  
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>  
F:1108-1133/Domain: transmembrane #status predicted <TM>  
F:1134-1163/Domain: intracellular #status predicted <INT>  
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 59.0%; Score 3513; DB 1; Length 1163;  
Best Local Similarity 61.4%; Pred. No. 1,1e-235;  
Matches 701; Conservative 138; Mismatches 296; Indels 6; Gaps 4;

QY 5 VLLTALTLCHGFNLDENAMTFQENARGFGSVVQLQSRVVGAPQEIIVAAANGSLY 64  
DB LLLTALATSLGFLNLTDELTAFVDSAGFGSVVGVVANSVWVGAPQKITAANGTGLY 67  
QY 65 OQDYSTGSCERLQVPEAVNMSGLSLAATTSPLLACGPTVHOTCSENTYVKGCLP 124  
DB 68 QCGYSTGACEPIGLQVPEAVNMSGLSLASTTSPLLACGPTVHCEGRNMYLTGLCP 127  
QY 125 LFGSNLRQPKPPALRGCPQEDSDIAPLIDGSGSIIPHDFRMEKEFVSTVMEQLKSK 184  
DB 128 LLGPT-  
QY 185 TLPSLMQYSEPRHFTKEFQNPNSRLVKPITQLLGRHTATGIRKRVRELPNYNG 244  
DB 186 TOFLMQPSNRKPTHTPEEFRTNSPLSLASVHQLQGFTYTAIQNVVHRLFHASYG 245  
QY 245 ARKNAFKILVITDCEKFGDPLGVEDVPEADREGVIRYVIGVDGDAFSEKSRQELATIA 304  
DB 246 ARRDATKILVITDCKEGSDLDYKDVIPMDADAAGIIRYALGVGLAFQNRNSWRELNDIA 305  
QY 305 SKPRDRHVQNNFALYTIQNLREKIPALHGTGTGSSSPHEHMSQEGSAITSNGP 364  
DB 306 SKPSQEHFHKVEDFDALXIOQLKEKIPALHGTGTGSSSPHEHMSQEGSAITSNGP 365  
QY 365 LLSTVGSVDWAGGVLYTSKEKSTFINTRVDSMDNDAYLVGAALILRNVRQSLVLGAP 424  
DB 366 VLGVGSFTWGGAPLYPNKSPFINNSQENVDRDSVLYSTELALWGVQSLVLGAP 425  
QY 425 RYQHIGLVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYE 484  
DB 426 RYQHTGKAVITQVSRQWRMAEVTGTQIGSYFGASLCSVDVDTGSTDVLIGAPHYE 485  
QY 485 QTRGGVSVCLPRGORAWCDVLYGEOQPCWCFGAALTVLGDVNGDKLTDTVALGAP 544  
DB 486 QTRGGVSVCLPRGWR-RWNCDAVLYGEOQPCWCFGAALTVLGDVNGDKLTDTVALGAP 544  
QY 545 GEENRGAVYLFHGTSGSGISPSHSQRTAGSLSPRLQYFGQSLSGGQDLTQDGLVDLAV 604  
DB 545 GEENRGAVYLFHGTSGSGISPSHSQRTAGSLSPRLQYFGQSLSGGQDLTQDGLVDLAV 604  
QY 605 GAQGHVLLRSQPLRVKVAIMEENPREVARNVPCNDGVVKEAGEVRVCLHVQKSTRD 664  
DB 605 GARGOVLRLTRPVLWUGVSHQFTPAEIPRPAFCEQVSEQLTVGNSNICLIDKSKNS 664  
QY 665 FLREGIQSVVYTDIALDGRPHGRAVNETKNSRTRQVTLGTCETLKLQLPNCIE 724  
DB 665 LLGSRDLQSSVTLDALDPGRSLSPRATFOETKRSLSRVRLGLKACENFNLPLPSQVE 724  
QY 725 DPVSPVLRNFSLVGTPLSFGNLRPVLAEDAQELTALPFPKNCQNDNICODLSIT 784  
DB 725 DSVTPITURLNFTLVGKPLLRNPLMLAADAQRYFTASLPFKNCQNDNICODLSIT 784  
QY 785 FSEMSLCLVVGGRPREFNVTVTRNDGEDSVYRTQVTFPPFLDLSYRVKYSTLQNSORSW 844  
DB 785 FSPFGLKSLVGSNLELNAEVMWVNDGEDSVYRTQVTFPPFLDLSYRVKYSTLQNSORSW 844  
QY 845 RLACESASTSVGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904  
DB 845 HLTCDSPVPG--SQGTWSTSCRIHILPRGQAQITFLATFDVSPKAVLGDRLILTANVSS 902  
QY 905 ENNMPTNKTEFQLELPVKYAVVMVTVTSHGVSTVLYNFTAS-ENTSRVMHQYQVSNLQ 963

DB 903 ENMPTNKTEFQLELPVKYAVVMVTVTSHGVSTVLYNFTAS-ENTSRVMHQYQVSNLQ 962  
QY 964 RSLPISVLVLPVRLNQTWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVVC 1023  
DB 963 RDLFVSINFVPELNCQEAVMQVMDVEVHPQNPFLSRCSSEKIAPPASDFLAHIQNPVLDC 1022  
QY 1024 SIACVQRIQCDIFFFGIOEENFATLKGNSLDFDAVYIKTSHNLLLVSTAELFNDVSFTLL 1083  
DB 1023 SIACVQRIQCDIFFFGIOEENFATLKGNSLDFDAVYIKTSHNLLLVSTAELFNDVSFTLL 1082  
QY 1084 PGQAFVRSOTETKVEPFEVNPFLIVGSSVGGLLLLALITAAALYKLGFPKQYKDMMS 1143  
DB 1083 PGQAFVRSOTETKVEPFEVNPFLIVGSSVGGLLLLALITAAALYKLGFPKQYKDMMS 1142  
QY 1144 E 1144  
DB 1143 E 1143

RESULT 4  
S03308  
N:Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte functi  
C:Species: Homo sapiens (man)  
C:Date: 28-Feb-1990 #sequence, revision 28-Feb-1990 #text\_change 20-Aug-1999  
C:Accession: S03308; A47458; A47565; A48759; S36044  
J: Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.  
J: Cell Biol. 108, 703-712, 1989  
A:Title: Primary structure of the leukocyte function-associated molecule-1 alpha subun  
A:Reference number: S03308; MUID:89139587; PMID:2537322  
A:Accession: S03308  
A:Molecule type: mRNA  
A:Residues: 1-1170 <LAR>  
A:Cross-references: EMBL:Y00796; NID:g31421; PIDN:CAA68747.1; PID:g31422  
A:Note: part of this sequence was confirmed by protein sequencing  
R: Cornwell, R.D.; Gollahan, K.A.; Hickstein, D.D.  
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993  
A:Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) p  
A:Reference number: A47458; MUID:93248261; PMID:8097887  
A:Accession: A47458  
A:Molecule type: DNA  
A:Residues: 1-20 <COR>  
A:Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBI:130863)  
R: Shelley, C.S.; Farkhad, O.C.; Arnaout, M.A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 5384-5386, 1993  
A:Title: Identification of cell-specific and developmentally regulated nuclear factors  
A:Reference number: A47565; MUID:93281759; PMID:8099450  
A:Accession: A47565  
A:Molecule type: DNA  
A:Residues: 1-20 <SHE>  
A:Cross-references: GB:M95609  
R: Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.  
J. Biol. Chem. 268, 19305-19311, 1993  
A:Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.  
A:Reference number: A48759; MUID:93374910; PMID:8103515  
A:Accession: A48759  
A:Molecule type: DNA  
A:Residues: 1-20 <NUE>  
A:Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406  
C:Genetics:  
A:Gene: GDB:ITGAL; CD11A  
A:Cross-references: GDB:119757; OMIM:153370  
A:Map position: 16p11.2-16p11.2  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hor  
C:Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; ti  
P:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted  
F:154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.2%; Score 1559; DB 2; Length 1170;  
Best Local Similarity 34.3%; Pred. No. 9.5e-100;  
Matches 410; Conservative 210; Mismatches 464; Indels 110; Gaps 38;



A>Note: part of this sequence was confirmed by protein sequencing  
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.  
J. Immunol. 150, 480-490, 1993  
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-1  
n during evolution.  
A;Reference number: A46526; MUID:93123748; PMID:8419480  
A;Accession: A46526  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-499,501-1153 <FILE>  
A;Cross-references: GB:S52227; NID:9263047; PIDN:AAB24821.1; PID:9263049  
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature  
A;Note: sequence extracted from NCBI backbone (NCBI:P121963)  
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.  
Biochim. Biophys. Acta 874, 368-371, 1986  
A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp  
A;Reference number: A90664; MUID:87076671; PMID:3539202  
A;Accession: A26091  
A;Molecule type: protein  
A;Residues: 17-31 <PIE>  
A;Experimental source: granulocytes  
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.  
Blood 79, 865-870, 1992  
A;Title: Characterization of the myeloid-specific CD11b promoter.  
A;Reference number: I52567; MUID:92144986; PMID:1346576  
A;Accession: I52567  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-9 <RES>  
A;Cross-references: GB:M84477; NID:9180184; PIDN:AAA51960.1; PID:9553219  
A;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
C;Genetics:  
A;Gene: GDB:ITGAM; CRJA  
A;Cross-references: GDB:120599; OMIM:120980  
A;Map position: 16p11.2-16p11.2  
A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>  
F;17-1108/Domain: extracellular #status predicted <EXT>  
F;148-318/Domain: von Willebrand factor type A repeat homology <WA2>  
F;465-478/Region: calcium/magnesium binding #status predicted  
F;530-538/Region: calcium/magnesium binding #status predicted  
F;593-601/Region: calcium/magnesium binding #status predicted  
F;1109-1134/Domain: transmembrane #status predicted <TM>  
F;1135-1153/Domain: intracellular #status predicted <INT>  
P;86.240.391.469.693.697.735.802.881.901.912.941.947.979.994.1022.1045.1051.1076/Binding

Query Match 100.0%; Score 5956; DB 1; Length 1153;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCHGPNLDNENAMTFQENARGFGQSVVUQSSRVVVGAPQEIIVANQR 60  
DB 1 MALRVLLLTALTLCHGPNLDNENAMTFQENARGFGQSVVUQSSRVVVGAPQEIIVANQR 60

QY 61 GSLVQCDYSTGSCPEIRIQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120  
DB 61 GSLVQCDYSTGSCPEIRIQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120

QY 121 GLCFLFGSNLRQOPKPFQALRGCPQEDSDIAPLDGSGSIIPHDPRMKFEVSTWEOQL 180  
DB 121 GLCFLFGSNLRQOPKPFQALRGCPQEDSDIAPLDGSGSIIPHDPRMKFEVSTWEOQL 180

QY 181 KKSXTLFLSMQYSEEFRIHTFKFQFNPNPRSLVKPTITQLLGRTHATGIRKVVRLFN 240  
DB 181 KKSXTLFLSMQYSEEFRIHTFKFQFNPNPRSLVKPTITQLLGRTHATGIRKVVRLFN 240

QY 241 ITNGARKNAFKILVITDGEKFGPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300  
DB 241 ITNGARKNAFKILVITDGEKFGPLGYEDVIPADREGVIRYVIGVDAPRSEKSRQEL 300

QY 301 NTIASKPRDHVFOVNNFEALKTIONLRBKI FAIEGTQTGSSSSFEHMSQEGFSAIT 360  
DB 301 NTIASKPRDHVFOVNNFEALKTIONLRBKI FAIEGTQTGSSSSFEHMSQEGFSAIT 360

QY 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLV 420  
DB 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLV 420

QY 421 LGAPRYOHIGLVAMFRONTGWESNANVKGTQICAYFGASLCSDVDNSGSTDILVLIGAP 480  
DB 421 LGAPRYOHIGLVAMFRONTGWESNANVKGTQICAYFGASLCSDVDNSGSTDILVLIGAP 480

QY 481 HYYOTRGQGVSVCPPLPRGQARWQCDVLYGEGQPGWRFGAALTIVLGVNGDKLTDVA 540  
DB 481 HYYOTRGQGVSVCPPLPRGQARWQCDVLYGEGQPGWRFGAALTIVLGVNGDKLTDVA 540

QY 541 IGAPGEDNNGAVLPHGTSGSGLSPSHSORIAGSKLSPLOYGOSLSGQDLMQGLV 600  
DB 541 IGAPGEDNNGAVLPHGTSGSGLSPSHSORIAGSKLSPLOYGOSLSGQDLMQGLV 600

QY 601 DLTVGAQGHVLLLSQPLVAVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQK 660  
DB 601 DLTVGAQGHVLLLSQPLVAVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQK 660

QY 661 STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKSTRTROTQVTLGTTCETLKLQLP 720  
DB 661 STRDLREGQIOSVVTYDLALDSGRPHSRVAFNETKSTRTROTQVTLGTTCETLKLQLP 720

QY 721 NCIEDPVSPIVLRNLSLVGTPLSAFGNLRPVLAEQAORLFTALFPPEKNCGNNDICODD 780  
DB 721 NCIEDPVSPIVLRNLSLVGTPLSAFGNLRPVLAEQAORLFTALFPPEKNCGNNDICODD 780

QY 781 LSITFSFMSLDCLVVGCPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSR 840  
DB 781 LSITFSFMSLDCLVVGCPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSR 840

QY 841 QRSNRLACESASSTEVSGALKSTSCSINHIFIPFENSEVTENITFDVDSKASLGKLLKA 900  
DB 841 QRSNRLACESASSTEVSGALKSTSCSINHIFIPFENSEVTENITFDVDSKASLGKLLKA 900

QY 901 NVTSENMPRTNKTPEQLPVLKAVVMVTSHGVSSTKYLNFTASENTSRVMHQYQVSN 960  
DB 901 NVTSENMPRTNKTPEQLPVLKAVVMVTSHGVSSTKYLNFTASENTSRVMHQYQVSN 960

QY 961 LGQRSLEPISLVLPVVRNLTWDRPQVTFNSNLSTCHTKERLPSHSDFLAELKAPV 1020  
DB 961 LGQRSLEPISLVLPVVRNLTWDRPQVTFNSNLSTCHTKERLPSHSDFLAELKAPV 1020

QY 1021 VNCIAVCQRIQCDIPFGIOBENATLKNLSFDWYIKTSHNHLIIVSTABILFNDSVF 1080  
DB 1021 VNCIAVCQRIQCDIPFGIOBENATLKNLSFDWYIKTSHNHLIIVSTABILFNDSVF 1080

QY 1081 TLLPGQAGPVRSOTETKVEPFEVNPPLPLIVGSSVGLLALLIITAALYKLGFFKQYKD 1140  
DB 1081 TLLPGQAGPVRSOTETKVEPFEVNPPLPLIVGSSVGLLALLIITAALYKLGFFKQYKD 1140

QY 1141 MMSEGGPPGAEPPQ 1153  
DB 1141 MMSEGGPPGAEPPQ 1153

RESULT 2  
S00551  
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse  
N;Alternate names: complement-3 receptor alpha chain  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 22-Oct-1999  
C;Accession: S00551; I59078  
R;Pyrcel, R.  
EMBO J. 7, 1371-1378, 1988  
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the  
A;Reference number: S00551; MUID:88312584; PMID:3044779  
A;Accession: S00551

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:05:16 ; Search time 14.7639 Seconds  
(without alignments)  
7512.163 Million cell updates/sec

Title: US-09-902-481b-1

Perfect score: 5956

Sequence: 1 MALRVLLTALTLCHGFNLD.....FTKQYKDMSEGGPPGAEPQ 1153

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5956	100.0	1153	1 RWHULB	cell surface glyco
2	4545	76.3	1153	2 S00551	leukocyte surface
3	3513	59.0	1153	1 RWHULC	cell surface glyco
4	1559	26.2	1170	2 S03308	cell surface glyco
5	1549	26.0	1163	2 I56126	lymphocyte function
6	1163	19.5	1179	2 A53213	integrin alpha-E c
7	1108.5	18.6	1151	2 A45226	integrin alpha-1 c
8	1099	18.5	1170	2 I45314	integrin alpha 2 s
9	1089	18.3	1178	2 S44142	VIa-2 protein homo
10	1087.5	18.3	1181	2 A33998	integrin alpha-2 c
11	1082	18.2	1180	2 A35854	integrin alpha-1 c
12	667	11.2	1039	2 A41131	lymphocyte-Peyer's
13	638	10.7	1038	2 S06046	integrin alpha-4 c
14	630	10.6	1035	2 I58409	integrin alpha-9 c
15	614.5	10.3	1041	2 T31437	integrin alpha cha
16	579.5	9.7	1042	2 JC7294	alnap integrin -
17	572.5	9.6	1051	2 A35761	cell surface glyco
18	567.5	9.5	1053	2 I55534	VIa-3 alpha subuni
19	555.5	9.3	1034	2 S44250	integrin alpha-5 c
20	550	9.2	1034	2 A36108	integrin alpha-v c
21	535	9.0	1044	2 T10050	integrin alpha-v c
22	532	8.9	1049	2 A27079	fibronectin recept
23	532	8.9	1073	2 B36429	integrin alpha-6 c
24	530.5	8.9	1072	2 A48457	integrin alpha-6 c
25	529.5	8.9	1051	2 A40021	integrin alpha-3 c
26	526	8.8	1048	2 A27421	integrin alpha-5 c
27	525.5	8.8	1091	2 A41543	integrin alpha-6 c
28	517	8.7	1044	2 S16516	integrin alpha-8 c
29	505	8.5	1394	2 A29637	position-specific

RESULT 1  
RWHULB  
cell surface glycoprotein CD11b precursor [validated] - human  
N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein 1  
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000  
C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567  
R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.  
J Biol. Chem. 263, 12403-12411, 1988  
A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C  
B.  
A;Reference number: A31108; MUID:88315033; PMID:2457584  
A;Accession: A31108  
A;Molecule type: mRNA  
A;Residues: 1-1153 <COR>  
A;Cross-references: GB:J03925; NID:gi87284; PIDN:AAA59544.1; PID:G307148  
A;Note: part of this sequence was confirmed by protein sequencing  
R;Arnaut, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.  
J. Cell Biol. 106, 2153-2158, 1988  
A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor  
A;Reference number: A28915; MUID:88257215; PMID:2454931  
A;Accession: A28915  
A;Molecule type: mRNA  
A;Residues: 1-499,501-965,'P',967-1153 <ARN>  
A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAA5  
A;Note: the authors translated the codon TAC for residue 1129 as Thr  
A;Note: part of this sequence, including the amino end of the mature protein, was confi  
R;Shelley, C.S.; Arnaut, M.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991  
A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally re  
A;Reference number: A41600; MUID:92073318; PMID:1683702  
A;Accession: A41600  
A;Molecule type: DNA  
A;Residues: 1-9 <SHE>  
A;Cross-references: GB:M76724; NID:gl80018; PIDN:AAA58410.1; PID:q553215  
R;Arnaut, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.  
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988  
A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe  
A;Reference number: A34193; MUID:86190151; PMID:2833753  
A;Accession: A30892  
A;Molecule type: mRNA  
A;Residues: 917-1042 <AR2>  
A;Cross-references: GB:M18044  
R;Hickstein, D.D.; Hickey, W.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.  
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989  
A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept  
A;Reference number: A32218; MUID:89098893; PMID:2563162  
A;Accession: A32218  
A;Molecule type: mRNA  
A;Residues: 9-1153 <HIC>  
A;Cross-references: GB:J04145; NID:gl89068; PIDN:AAA59903.1; PID:q386975

30 496.5 8.3 1146 2 S40311 integrin - fruit f  
31 495.5 8.3 1039 2 A34269 integrin alpha-2b  
32 492.5 8.3 1037 2 A60163 glycoprotein Iib -  
33 490 8.2 1137 2 JC5950 integrin alpha-7 c  
34 488 8.2 1135 2 I61186 alpha-7 integrin -  
35 486 8.2 126 2 B30892 leukocyte adhesion  
36 474.5 8.0 1226 2 S44824 p5472.1 protein -  
37 469.5 7.9 1106 2 S38783 integrin alpha cha  
38 454.5 7.6 1139 2 S28277 hypothetical prote  
39 450 7.2 1115 2 T09433 integrin alpha v c  
40 430.5 7.1 1115 2 T09403 integrin alpha cha  
41 423.5 6.6 764 2 I36916 glycoprotein Iib -  
42 391 5.2 1086 2 T18523 integrin alpha cha  
43 309 5.1 272 2 A55348 integrin alpha-1 -  
44 304.5 5.0 604 2 I36917 glycoprotein Iib -  
45 299 5.0 604 2 I36917

APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. US2003007278A1el Human 2  
FILE REFERENCE: 27866/35004  
CURRENT APPLICATION NUMBER: US/09/891,943  
CURRENT FILING DATE: 2001-06-26  
PRIOR APPLICATION NUMBER: 09/193,043  
PRIOR FILING DATE: 1998-11-16  
PRIOR APPLICATION NUMBER: 08/286,889  
PRIOR FILING DATE: 1994-08-05  
PRIOR APPLICATION NUMBER: 08/362,652  
PRIOR FILING DATE: 1994-12-21  
PRIOR APPLICATION NUMBER: 08/943,363  
PRIOR FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1161  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-891-943-2

Query Match 57.6%; Score 3388; DB 10; Length 1161;  
Best Local Similarity 58.9%; Pred. No. 5e-304;  
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

Qy 1 ENLDTENAMTQENARGGQSVVOLQGSVVVGAPOEIVAAQNGSLYQCDYSTGSCBPI 60  
Db 17 ENLDEEETIPEQDAGGQSVVOLQGSVVVGAPOEIVAAQNGSLYQCDYSTGSCBPI 76

Qy 61 RLQVPEAVNNLSGLSLAATSPPOLLAGCTVTHQTCSENTYVVKGLFPLGSLNRQQPOK 120  
Db 77 PLHTRPRAVNNLSGLTLAASNGSLLACGTLTHVCGENSYSKGSCLLGSRW-ELIQT 135

Qy 121 FPEALRGCPQEDSIAFLIDSGSII PHDPRMKELVSTIMEQLKXKTLFSLMOYSEEF 180  
Db 136 VPDAPPECPHQMEDIIVFLIDSGSIDQNDQMKGFVQAVMGQPEGDTLFPALMQYSNLL 195

Qy 181 RIHFPEKFNPNPRSLIKPITOLLGRTHTATGLRVRELFTNITGARKNAFKILFL 240  
Db 196 KIHFTFTQFRTSPSQSLVDIVQLKGLTFTATGILTVVTLFHHKNGARKSAKILIVI 255

Qy 241 TDGEKFGDPLGYEDVIELDREGVIRVYLGFDAPRSEKSOBLNTVASKPPDRDHVQAN 300  
Db 256 TDGOKYKDPLEYSVDVIPAQERAGIIRVAGVGHAFQGPARTARQELNTISSAPPODHVEKD 315

Qy 301 NFEALKTVQNLREKI PAIEGTOTGSSSSFEHMSQGFSAITNSGELLSTVGSYDWAG 360  
Db 316 NPAALGSIQKQKQEKIYAVEGTQGRASSSFHMSQGFSTALTMGDLFGAVGSFWSG 375

Qy 361 QVFLYTSKEKSTFINMTRVDSMDNDVILGYAAAIILRRNVQSLVGLAPRYOHIGLVAMFR 420  
Db 376 GAFLYPPNMSPTFINMSQENVMDRDSVLYGTELALWKGQNLVGLAPRYOHTGKAVFT 435

Qy 421 QNTGWESNANVKCTOICAVGASLCSVDVDSNGSTDLVILGAPHYVEOTRGQSVCP 480  
Db 436 QVSGQWRKAEVTGTQGSYFGASLCSVDVDSNGSTDLVILGAPHYVEOTRGQSVCP 495

Qy 481 PRGORARQCDVILYGEQGPWGRFGAALTVLGVNGDKLTDVAIGAPGBEDNRGAVYLF 540  
Db 496 PRGORVQWQCDVILRGEGHPWGRFGAALTVLGVNEDKLDVAIGAPGBEDNRGAVYLF 555

Qy 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSCGQDLTMDGLVDITVGAQGHVILLRSQ 600  
Db 556 HGASESGISPSHSQRIASSQLSPRLQYFGQSLSCGQDLTMDGLVDITVGAQGHVILLRS 615

Qy 601 PVLKVKAIMFNPFEVARNVPECNDQVVKGEAGEVVRVCLHVQKSTRDLRREGQIQSVVT 660  
Db 616 PVLKVGVMRFPVEVAVAKVVRWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy 661 YDLALDSGRPHSRVAFNETKNSRQTVGLTQTCBTLKQLPNCIEDPVSPIVLELNF 720  
Db 674 FDLALDGRLTSAIFNETKNPTLTRKTLGLGHCETKLLLPDCVEDVVSPIILHNF 733

Qy 721 SLVGTPLSARGNLRPVLAEDAQRLEFTALPFEKNCNDNICODDLSTTSFMSLDCLVVG 780  
Db 734 SLVREPIPSQNLRPVLAVGSQLFTRASLEFERNCCQDGLCBGDLGVTLSFSGLOTITVG 793

Qy 781 GPREFNVTVTRNDGEDSYETQVTFPPLDLSYRKVSTLQNRQSRQSRWRLACASASSTEV 840  
Db 794 SSLELNVITVMNAGEDSYGTWVSLYYPAGLSHRVSVGAQKQPHQSALRLACETV-PTED 852

Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKXANTVSENMPRTNKTFF 900  
Db 853 EG-LRSSRCVSNRPIFHEGNGTFIVTFDVSYNATLGDRLMKASASSENKASSSKATF 911

Qy 901 QLELPVKIAYVMVTSHGVSSTKYNLF-TASENTSRVMQHOYQVSNLQSRSLPISLVFLVP 959  
Db 912 QLELPVKIAYVMVTSRQESTKYNFATSDKXKKEAHEHYRVNNLSQDLAISINFWVP 971

Qy 960 VRLNQTVIMDRPQVTFSEMLSSCTHYKERLPSSHDELAEKRPVNVNCSIAVQRIQCDI 1019  
Db 972 VLLNGVAVDVMVMEAFSQSL--PCVSEKPPQSHSDFLTQISRSPMLDCSIADCLQPRCDV 1029

Qy 1020 PFFGIQEEFNATLKGNSLFDWYIKTSHNHLIIYSTABILFNDSVFTLLPQGGAFVRSQTE 1079  
Db 1030 PSFSVQBELDFTLKGNSLFCGVVRETIQKVLVVSVAHITEDTSVYSQLPQGEAFMRAQME 1089

Qy 1080 TKVEPEVPNPLPIIVGSSVSGGIIILALITAAALYKLGFFKRYKMMSE 1128  
Db 1090 MVLEDEEVYNAIPIIMGSSVSGALLLALITATUYKLGFFKRYKEMLEED 1138

Search completed: June 7, 2004, 17:38:44  
Job time : 42.8905 secs

QY 721 SLVGTPLSAGNLRVLAEDAOQLPTALPPEKNCNGNDNICODDLSITPSPMSLDCLVVG 780  
 DB 737 TLVGKPLLAFLNLRPMLAALQRIFTASLPPEKNCNGADHICODMIGISFSPGUKSLVG 796  
 QY 781 GPREFNVTVRNDEDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLACESASTEV 840  
 DB 797 SNLENAEVMWMDGSDSYGTTTFESHAPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854  
 QY 841 SGALKSTSCSNHDPFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900  
 DB 855 SQGTWSTSCRINHLIFRGAGITLAFDVSFKAVGLDRLLLIANVSSENNIPRTSTIF 914  
 QY 901 QLELPVYAVVYVTSYSHGVSTKYLNFTAS-ENTSRVMOHQVQSNLQORSIPISLVLVP 959  
 DB 915 QLELPVYAVVYVSSHEQFTKYLNFSESEKSHVAMHRVQVNNLQORDLPVSNWVP 974  
 QY 960 VRLNQTWIDPEQVTFSENLSSTCHTVERLPDPSHDFLAELKAPVNCSTAVOCRIODI 1019  
 DB 975 VELNOEAVMDVESHQNPNSLRCSSEKIAPPASDFLAHTQKNPVLDCSIAGCLRFCDV 1034  
 QY 1020 PFGIOEFNATLKNLSFDFWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSOTE 1079  
 DB 1035 PSFSVQEBELDTLKNLSFGWVROILOKKYSVVSVAELIIPDTSYVSOLPGQAFVRSOTE 1094  
 QY 1080 TKVPEFVPEPLPIVGVSSVGLLLALITALYKLGFFKRYQKDMNSE 1128  
 DB 1095 TVLEKYVHNPIPLIVGSSIGGLLLALITALYKLVGVGFFKRYQKDMNSE 1143

RESULT 14

US-09-350-259-2  
 ; Sequence 2, Application US/09350259  
 ; Patent No. US20020062008A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.  
 ; APPLICANT: Van der Vliet, Monica  
 ; TITLE OF INVENTION: No. US20020062008A1 Human 2  
 ; FILE REFERENCE: 27866/35004  
 ; CURRENT APPLICATION NUMBER: US/09/350,259  
 ; CURRENT FILING DATE: 1999-07-08  
 ; EARLIER APPLICATION NUMBER: 09/193,043  
 ; EARLIER FILING DATE: 1998-11-16  
 ; EARLIER APPLICATION NUMBER: 08/173,497  
 ; EARLIER FILING DATE: 1993-12-23  
 ; EARLIER APPLICATION NUMBER: 08/286,889  
 ; EARLIER FILING DATE: 1994-08-05  
 ; EARLIER APPLICATION NUMBER: 08/362,652  
 ; EARLIER FILING DATE: 1994-12-21  
 ; EARLIER APPLICATION NUMBER: 08/943,363  
 ; EARLIER FILING DATE: 1997-10-03  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1161  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-350-259-2

Query Match 57.6%; Score 3388; DB 9; Length 1161;  
 Best Local Similarity 58.9%; Pred. No. 5e-304;  
 Matches 665; Conservative 166; Mismatches 288; Indels 8; Gaps 6;  
 QY 1 FNLDTENAMTFQENARFGQSVQLOQSRVVGAPQIVAAQNRGSLYQCDYGTGSCPI 60  
 DB 17 PNLDEBPTTFQEDAGFGQSVQVQFSGSLVVGAPLEVAAQNRGSLYQCDYGTGSCPI 76  
 QY 61 RLQVFEAVNMSGLSLAATSPOLLACGPTVHTQTCSENTYKGLCFLFGSNLRQOPQK 120  
 DB 77 PLHIRPEAVNMSGLTLAATSTNGSRLACGPTLHRCVGENSYKSGSLGSLGSRW-ELIQT 135  
 QY 121 PPRALRCPODSDIAPLIDGSSIIPEHDFRMEKLVSTIMEOLKSKTLFSLMOYSEEP 180

DB 136 VPDATECPEQEMDIIVPLIDGSSIQDNDPQNMKGFPVQAVMGOFECTDTLIPALMOYSNLL 195  
 QY 181 RIHFTYKBFONNPNRSLIKPTTOLLGRTHETATGLKRYVRELPNITNGARKAKFILL 240  
 DB 196 KIHFTTQRTSPSOQSLVDPIVOLKGLTFTATGILTVVTLPHKNGARKAKKILIVI 255  
 QY 241 TDCXKGDPLGYBDVIPELDREGVIRYVLGFGDAPSEKSRQBELNLTVASPPRDPHVFQAN 300  
 DB 256 TDQCKYKDPLEYSVDVPOAEKAGIRYAVGVHAFQGPATQBELNTISAPPOPHVFKVD 315  
 QY 301 NFALTKTVQONQARAKTIPATEGTQSSSPFHEHMSQEGSAAITSNGLPILLSTVSGVDWAG 360  
 DB 316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDELFLGAVGSPWSG 375  
 QY 361 GVFLYTKSKSTPINNTRVDSMDNDAYLGVAAAIILNRNVQSLVAGAPRYQHTIGLWAMER 420  
 DB 376 GAFLYPNWSPPIFNKSNQENVDWROSYLGYSTELALWKGVQNLVGLAPRYQHTGKAVFT 435  
 QY 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQOVSCPL 480  
 DB 436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQOVSCPL 495  
 QY 481 PRGORARWOCDAVLVGEQCPKGRFGAALTVLGDVNGDKLTQVAIGAPGEEDNRGAVYLF 540  
 DB 496 PRQVRVQWQCDVLRGEQHPGRFGAALTVLGDVNGDKLTQVAIGAPGEEDNRGAVYLF 555  
 QY 541 HGTSGSGISPSHESQRIAGSKLSPRLQYFGQSLSGGQDLTMDELVDLTGVAQGHVLLRSQ 600  
 DB 556 HGASBSGSIHESQRIAGSKLSPRLQYFGQSLSGGQDLTMDELVDLTGVAQGHVLLRSQ 615  
 QY 601 PVLKVKALMEFNPREVARNVPECDQVVKKEAGVYVCLHWQKSTDLREGQIOSVVT 660  
 DB 616 PVLKVGVMRFPSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSLDQ--GDIQSSVR 673  
 QY 661 YDLALDGPGRHRAVENETKSTRQTVLGTQTCETLKLQLPNCIEBPVSVIVRLNF 720  
 DB 674 FDLALDGPGRHRAVENETKSTRQTVLGTQTCETLKLQLPNCIEBPVSVIVRLNF 733  
 QY 721 SLVGTPLSAGNLRVLAEDAOQLPTALPPEKNCNGNDNICODDLSITPSPMSLDCLVVG 780  
 DB 734 SLVREPIPSQNLRLPVLAVGSDLPASLPPEKNCNGQDGLCBGDLGVTLSPSGLOTIVG 793  
 QY 781 GPREFNVTVRNDEDSYRTQVTFPPFLDLSYKVKSTLQNRQSRWRLACESASTEV 840  
 DB 794 SNLENAEVMWMDGSDSYGTTTFESHAPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854  
 QY 841 SGALKSTSCSNHDPFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEP 900  
 DB 853 EG-LRSSRCSVNHPIPHESGNGTFIVTFDVSVKATIGDRMLRASSENKASSKATP 911  
 QY 901 QLELPVYAVVYVTSYSHGVSTKYLNFTAS-ENTSRVMOHQVQSNLQORSIPISLVLVP 959  
 DB 912 QLELPVYAVVYVTSYSHGVSTKYLNFTAS-ENTSRVMOHQVQSNLQORSIPISLVLVP 971  
 QY 960 VRLNQTWIDPEQVTFSENLSSTCHTVERLPDPSHDFLAELKAPVNCSTAVOCRIODI 1019  
 DB 972 VLLNGVAVMDVESHQNPNSLRCSSEKIAPPASDFLAHTQKNPVLDCSIAGCLRFCDV 1029  
 QY 1020 PFGIOEFNATLKNLSFDFWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSOTE 1079  
 DB 1030 PSFSVQEBELDTLKNLSFGWVROILOKKYSVVSVAELIIPDTSYVSOLPGQAFVRSOTE 1089  
 QY 1080 TKVPEFVPEPLPIVGVSSVGLLLALITALYKLGFFKRYQKDMNSE 1128  
 DB 1090 MVLDEBVTNAPILMGSSVGLLLALITALYKLVGVGFFKRYQKDMNSE 1138

RESULT 15

US-09-891-943-2  
 ; Sequence 2, Application US/09891943  
 ; Publication No. US20030077278A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.

Query Match 18.6%; Score 1108.5; DB 1; Length 1151;  
 Best Local Similarity 28.0%; Pred. No. 1.1e-65;  
 Matches 346; Conservative 210; Mismatches 486; Indels 195; Gaps 44;

FT CARBOHYD 980 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1151 AA; 127637 MW; 5B3F3C1AABF52808 CRC64;

17 FNLDTENAMTFQENARG-FQGSVQVQ- --QGSRVVVGVAPOEIVAANRGSLVQCYDSTGS 72  
 1 FNDVENSMTFGPVEDMGYTVQYENEGKMWILGSLVGPQKRTGDKYKCPVGRGE 60  
 73 CEP-ILQVPVER- -----VNSLGLSLAATSPFOLLACGTVHQTCSNTYKGL 122  
 61 SLPCVKLDLFLVNTSIPNVTEVENMTFGSTL-VTPNPGFLACGPLVAYRGCHLHYTGI 119  
 123 CFLFGSNLQOQPKPPFALRGCPQSDSDFAPLIDGSGSIIIPDPRMKFVSTWTEQLK- 181  
 120 CSDVSFTFQVNSIAP- -VQEGSTQ-LDVIIVDGSNSIYPWD- -SVTAFNLNDLKRMDI 174  
 182 -KSKTLFSLMQSYSEFRIHFTKBFQNNPNRPSLVKPIQTOLGR-THTATGIRKVVRELF 239  
 175 GPQQTQGVQVQGENVTHFNLNKYSSTEEVLVAAKIVQRGQTMTALGTDARKEAP 234  
 240 NIINGARKNAFLVITDGKFGPLGVEDVVPADREGVIRVYVGVDAFR- ----SE 294  
 235 TEARGARGVKVMVITVDGESH-DNHRKJQVQDCENIQTFSIALGSYNGNLSTE 293  
 295 KSRQELNTIASPPRDHVPQVNNFEALKTIONQLREKIFAIEGTQTGSSSFHEMSQEG 354  
 294 KPVEEIKSIASEPTEKHFNNVDELALVTIVKTLGRIFALEATADQSAASFEMENSQTG 353  
 355 FSAATISNGLSLTSGSYDWAGCVFLYTSKE- ----KSTP-INTRVDSMDNDAYLYGAA 408  
 354 FSAHYSQDWMVLGAVGDMGTVVMQASQIIIPNFTFNVESTKKEPL-ASYLYGTV 412  
 409 AILLRNRVQSL-VLGAAPRYOHLGVAMPFRONTGMWESNANVKGTCIGAYFGACSCSYDVD 467  
 413 NSNTASSGDVLVIAGOPRYNHGTQVLIYVMEDGNIKILQTLGSGQISYFGSILTTTDD 472  
 468 SNGSTDLVLIGAPHY- -----YEQTR-QGQSVVCPPLRQGRARWCDAVLY 511  
 473 KDSNTDILLVGAFTWGTGEBQKQVYVVALNQTRFQWMSLEPIKQTCSSSRQHSCTT 532  
 512 GEGQOPWG-RFGAALTVLGVNGDKLTVAIGAPCEEDNRGANVLPFGTSGSIPSHSQ 570  
 533 ENKNEPGARFGTA-AAVKDLMDLGNFDNIVIGAPLEDHGGAVVIYHG-SOKTKIRKEYA 591  
 571 RIAGSKLSPLRYFQGSQSGGQBLTMDGLVDLTGVAQGHVLLRSQPVLRAKIMEFNPR 630  
 592 RIPSQGDGKTLKFFQSGINGEDLNGDGLTQVTIGLGGALFWSRDVAVKVTWTFEPN 651  
 631 EVARNVPEQNDQVVKKEAG- -EVRVCLHVQ-KSTRDLREGQIQSVVYTDLALDSRPH 687  
 652 KVNIOKQKCH- -MEGKETVCINATVCFEVLKSKEDTIYEADLQ- -YRVTLDSLRQI 704  
 688 SRAVFNET- ----KNSTERTQVGLCTCETLKLQLPNCIEPDPVSPVLRNLNFSLVGT 741  
 705 SRSPFGTQERKQVNRITVRKSEC- ----TKHSFMLDKHDFQDSVR- -ITLDFNLT-D 755  
 742 PLSAFNGPLVLAEDAQLFTALFPFEKXNCGNDIFQDDLSTITSFMSLCLVGVGPRE- 800  
 756 PENG- ----PVLDDSLPNSVHYEIPFAXDCGKKEKISDLSLRVATTEKOLLIVRSNDX 810  
 801 ENVTVTVRNGDGSYRTQVTFPPFLDLVSRKVVSTLQNRORSRELACESASSTEVSGAL 860  
 811 FVSLTVKTKDSANTRITVYSNLVFSGLERLQKD- ----SCSEN- - 853  
 861 KSTCSISNHPISPENSEVTFNITFDVDSKASLGN-KLLKLVANVTSNNMPTNKTEFQLE 919

854 HNITCKGVGPFLLRGEMVTFKILFQFNTSYLMENVTIYLSATSDESBPPETLSDNVVNIS 913  
 920 LPVKYAVWVWTSVSGVTKYLNFTASENTSRVMHQYQVSN- ----LGQRS- ----L 966  
 914 IPVKYEVGLQFVS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSPMPDEL 972  
 967 PLSIVP- ----LVPVLNQTVIWDREQVTFSENLSTCHTKE- ----RLPS 1007  
 973 KLSISFPNMTNGVPLVPTGLSS- ----SENANCRPHIDPFPSINSGKXMTT 1021  
 1008 HSDFLAELRKAPVNVNCSIAVCQRIQCCDIPFFGIOE- ----EFNATLK 1049  
 1022 STD- ----HLKRGITLDCNCKEATITCNLTSSDISQVNVSLILMKPTIKSVESLNLITIR 1078  
 1050 GNLSPDWIKTSHNELLIVTAEILFNDSVFTLLPGQAFVRSQSTETKVPFEPVNPPLP 1109  
 1079 GEL- ----RSENASVLSSN- ----QKRELATQISKDGLEGRVPL 1114  
 1110 -IVGSSVGGILLALITAAALYKLGFKRQYKDMWSE 1144  
 1115 WYLLSAPAGILLMLLILALWKIGFFRPLKXONEK 1151

RESULT 10  
 ITAH HUMAN  
 ID ITAH HUMAN STANDARD; PRT; 1189 AA.  
 AC QUKX5; QUKO1;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-11 precursor.  
 GN ITG11.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Petal heart, and Osteoblast;  
 RX MEDLINE=99417678; PubMed=10486209;  
 RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,  
 RA Wang S.-X., Morris C.M., Krissansen G.W.;  
 RT "Cloning, sequence analysis, and chromosomal localization of the novel  
 RT human integrin alphai1 subunit (ITG11).";  
 RL Genomics 60:179-187(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Petal muscle, and Uterus;  
 RX Medline=99395147; PubMed=10464311;  
 RA Welling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;  
 RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)  
 RT Integrin. A collagen-binding, i domain-containing, beta(1)-associated  
 RT integrin alpha-chain present in muscle tissues.";  
 RL J. Biol. Chem. 274:25735-25742(1999).  
 RN [3]  
 RP SEQUENCE OF 954-1188 FROM N.A.  
 RC TISSUE=Fibroblast;  
 RA Andreu N., Estivill x., Escarceller M., Sumoy L.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC - FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.  
 CC - SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11  
 CC ASSOCIATES WITH BETA-1.  
 CC - SUBCELLULAR LOCATION: Type I membrane protein.  
 CC - TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND  
 CC HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO  
 CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO  
 CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,  
 CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE, VERY LOW LEVELS IN  
 CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.  
 CC - DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING  
 CC FETAL MUSCLE CELLS (IN VITRO).  
 CC - DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.



Db 758 -----DHGXLDDGMPPTLLRVSPVFWNGCNEDEHCVDPDLVDARSPLTAMEYQORVLRKP 813  
Qy 782 -----SITFSNLSCLLVGGPREFWVTVANDGDSYRTQVTFPPDLDSYRKVSTILO 836  
Db 814 ADCCSAYLSFTTTFIESTQRVAVAEATLENGENAYSTVNLISQSANLQF--ASLIQ 871  
Qy 837 NORSORSWELACESASSTEVSGALKASTCSINHPFIPENSEVTFNITFDVDSKASLGNKL 896  
Db 872 KEDSDGS--IECVNEER-----RLQKQVCNVSYPFRAKAKVAFRLDFEF-SKSFILHHL 923  
Qy 897 LKANVTSENMPRTNKB--FQLELPKYAVTMVTSHGVSCTKY---LNFAS--ENTS 949  
Db 924 ETELAAGSDNEDRSTKEDNAPLRFHLKYEADVLFTRSSLSHSEVKNLSLERYDGG 983  
Qy 950 RVQHQYQVSNLGG--RSLPLSLVFLPVLNQTWIDRPOVTESEMLSSTC-----HTK 1002  
Db 984 PPSFCIFRIQNLGLPFIHGMKTIPTATRSNRLKLRDPLT-DEVANTSCNIGWST 1042  
Qy 1003 ERLPSHDFLABIRKAPVVCNCSIAVCQRIQCIPFGIOEBPNATLKGNLSPDY----- 1057  
Db 1043 EYRPTFVE--EDLRAPQLNHSNDSVSNICNIRLVP-NQEIHFLLGNL---WLSLKA 1096  
Qy 1058 IKTSHNHLIVGTAEILFNDVFTLLPGQCAFVRSQTEKVEPPEVN-----PLPLIV 1111  
Db 1097 LKXSKMKNWALQROFH-SFP-----IFREDPSRQIVFHSKQEDMQVDPIWIV 1147  
Qy 1112 GSGVGLLILLALITAAKYLGKFFK--ROYKMMSEGGPPGAP 1152  
Db 1148 GSTLGLLILLALLVLAWKLGFFRSARRR-----PGLDP 1183  
RESULT 11  
IT92\_BOVIN STANDARD; PRT; 1170 AA.  
AC P53710;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)  
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).  
GN IT92.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94193647; PubMed=7511592;  
RA Kamata T., Puzon W., Takada Y.;  
RT Identification of putative ligand binding sites within I domain of  
RL J. Biol. Chem. 269:9659-9663(1994).  
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND B-CADHERIN. IT  
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN  
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE  
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
CC EXTRACELLULAR MATRIX.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
CC ASSOCIATES WITH BETA-1.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L25886; AAB59255.1; -.  
DR PIR; I45914; I45914.  
DR HSSP; P17301; IACX.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR SMART; SM00191; Int alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS50234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.  
FT SIGNAL 1 18 POTENTIAL.  
FT CHAIN 19 1170 INTEGRIN ALPHA-2.  
FT DOMAIN 19 1121 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1122 1143 POTENTIAL.  
FT DOMAIN 1144 1170 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 34 92 FG-GAP 1.  
FT REPEAT ? ? FG-GAP 2.  
FT DOMAIN 177 367 VWFA.  
FT REPEAT ? ? FG-GAP 3.  
FT REPEAT 423 475 FG-GAP 4.  
FT REPEAT 477 538 FG-GAP 5.  
FT REPEAT 540 599 FG-GAP 6.  
FT REPEAT 604 656 FG-GAP 7.  
FT CA\_BIND 488 496 POTENTIAL.  
FT CA\_BIND 552 560 POTENTIAL.  
FT SITE 472 474 CELL ATTACHMENT SITE (POTENTIAL).  
FT SITE 1146 1150 GFFKR MOTIF.  
FT DISULFID 72 81 BY SIMILARITY.  
FT DISULFID 669 726 BY SIMILARITY.  
FT DISULFID 778 784 BY SIMILARITY.  
FT DISULFID 854 865 BY SIMILARITY.  
FT DISULFID 1008 1039 BY SIMILARITY.  
FT DISULFID 1044 1049 BY SIMILARITY.  
FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 688 688 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 945 945 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1063 1063 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1070 1070 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARIANT 580 580 G -> V.  
FT VARIANT 588 588 R -> K.  
FT VARIANT 725 725 R -> S.  
SQ SEQUENCE 1170 AA; 128929 MW; BECEFC5F2449FB1 CRC64;

Query Match 18.5%; Score 1099; DB 1; Length 1170;

Best Local Similarity 27.8%; Pred. No. 5e-65;  
Matches 340; Conservative 215; Mismatches 498; Indels 168; Gaps 47;

Qy 11 LTLCHGNLDTENAMTFQ-ENARFGQSVVQL-----OGSRVVVVGAPQSEIVAAORGLSYQC 66  
Db 13 LNCVVAYVGLPKAKIFSGPSSEQGVAYVQGFIPKGNWLLVGSFWSGPPKRNKRGVDVYKC 72  
Qy 67 --DYSTGSCPEIRIQ-----VPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNT 117  
Db 73 PVDLSITTCCKLNLTQSTSMNSVTMTKMTNLSGLTLTRNVGTGGLTGGPLWAOCCGSOY 132  
Qy 118 YVKGCLFLFGSNLRQOQPKTPEALRGCPQBDSDIAFLIDGSGSIIIPHDFRMKKEFVSTVM 177





RT possible collagen-binding domain.";  
 RL J. Cell Biol. 109:397-407 (1989).  
 RP [2]  
 RA Rieder M.J., Armet T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,  
 RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.  
 RA MEDLINE=98019223; PubMed=9353312;  
 RA Emley J., King S.L., Bergelson J.M., Liddington R.C.;  
 RL "Crystal structure of the I domain from integrin alpha2beta1.";  
 RT J. Biol. Chem. 272:28512-28517 (1997).  
 RN [4]  
 RP VARIANT HPA-5 (BR).  
 RA MEDLINE=94043762; PubMed=7901236;  
 RA Santos S., Kalb R., Walke M., Kiesel V., Mueller-Eckhardt C.,  
 RA Newman P.J.;  
 RT "The human platelet alloantigens Br(a) and Br(b) are associated with a  
 RT single amino acid polymorphism on glycoprotein Ia (integrin subunit  
 RT alpha 2)." ;  
 RL J. Clin. Invest. 92:2427-2432 (1993).  
 RN [5]  
 RP VARIANT GLU-534.  
 RA MEDLINE=20260009; PubMed=1074142;  
 RA Kroll H., Gardemann A., Fechter A., Haberbosch W., Santos S.;  
 RA "The impact of the glycoprotein Ia collagen receptor subunit A1648G  
 RT gene polymorphism on coronary artery disease and acute myocardial  
 RT infarction." ;  
 RL Thromb. Haemost. 83:392-396 (2000).  
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT  
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN  
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE  
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2  
 CC associates with beta-1. Interacts with HP95.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- POLYMORPHISM: Position 534 is associated with platelet-specific  
 CC alloantigen HPA-5 (BR). HPA-5A/BR(A) has Lys-534 and HPA-5B/BR(B)  
 CC has Glu-534. HPA-5B is involved in neonatal alloimmune  
 CC thrombocytopenia (NAIT or NAITP). The K534E polymorphism may play a  
 CC role in coronary artery disease (CAD).  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD49b entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X17033; CAA34894.1;  
 CC EMBL: AF512556; AM34795.1;  
 CC PIR: A33998; A33998.  
 CC PDB: 1AOK; 25-NOV-98.  
 CC PDB: 1DZ1; 02-AUG-01.  
 CC Genew: HGNC:6137; ITGA2.  
 CC MIM: 192974;  
 CC GO: GO:0008305; C:integrin complex; TAS.  
 CC GO: GO:0005886; C:plasma membrane; TAS.  
 CC GO: GO:0004895; F:cell adhesion receptor activity; TAS.  
 CC GO: GO:0005518; F:collagen binding; TAS.  
 CC GO: GO:0007596; F:blood coagulation; TAS.  
 DR GO: GO:0007160; P:cell-matrix adhesion; TAS.  
 DR GO: GO:0007397; P:histogenesis and organogenesis; TAS.  
 DR InterPro: IPR000413; Integrin\_alpha.  
 DR InterPro: IPR002035; VWF\_A.  
 DR Pfam: PF01839; FG-GAP 3.  
 DR Pfam: PF00357; Integrin\_A; 1.  
 DR Pfam: PF00092; vwa; 1.  
 DR SMART: SM00191; Int\_alpha; 5.  
 DR SMART: SM00327; VWF; 1.  
 DR PROSITE: PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE: PS00234; VWFA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;  
 KW 3D-structure.  
 FT SIGNAL 1 29  
 FT CHAIN 30 1181 INTEGRIN ALPHA-2  
 FT DOMAIN 30 1132 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1133 1154 POTENTIAL.  
 FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1155 1161 INTERACTION WITH HP95.  
 FT REPEAT 45 103 FG-GAP 1.  
 FT REPEAT ? ? FG-GAP 2.  
 FT DOMAIN 188 378 VWFA.  
 FT REPEAT 378 433 FG-GAP 3.  
 FT REPEAT 434 486 FG-GAP 4.  
 FT REPEAT 488 549 FG-GAP 5.  
 FT REPEAT 551 610 FG-GAP 6.  
 FT REPEAT 615 667 FG-GAP 7.  
 FT CA BIND 499 507 POTENTIAL.  
 FT CA BIND 563 571 POTENTIAL.  
 FT CA BIND 627 635 POTENTIAL.  
 FT SITE 1157 1161 GREY MOTIF.  
 FT DISULFID 83 92 BY SIMILARITY.  
 FT DISULFID 680 737 BY SIMILARITY.  
 FT DISULFID 789 795 BY SIMILARITY.  
 FT DISULFID 865 876 BY SIMILARITY.  
 FT DISULFID 1019 1050 BY SIMILARITY.  
 FT DISULFID 1055 1060 BY SIMILARITY.  
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 699 699 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 534 K -> E (IN ALLOANTIGEN HPA-5B;  
 FT dbSNP:1801106).  
 FT /FTID=VAR\_003977.  
 FT TURN 170 171  
 FT STRAND 173 180  
 FT TURN 183 184  
 FT HELIX 188 199  
 FT TURN 200 201  
 FT STRAND 204 204  
 FT TURN 206 207  
 FT STRAND 209 216  
 FT STRAND 220 224  
 FT TURN 226 228  
 FT HELIX 232 240  
 FT TURN 241 241  
 FT HELIX 252 262  
 FT TURN 263 264  
 FT HELIX 266 268  
 FT TURN 269 269  
 FT STRAND 275 282  
 FT HELIX 289 291  
 FT TURN 292 301  
 FT TURN 302 303  
 FT STRAND 304 311  
 FT HELIX 313 317







Db 311 GHYNRGNLSTKVFBEIKSIASRPTKHPFNVSDELALVTIVKALGERIFALEATADQSA 370  
Qy 344 SSEHEMSQGFSAATSNPILLSVGSYDMAGVFLYTSKEKSTFINMT--RVDSMDND 401  
Db 371 ASFEHMSQGFSAHYSQDMVLMGAVGAYDNGTVVWQXANQVIPHNTTFFQTEPAQKNE 430  
Qy 402 ---AYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMFRQNTGMWESNANVKGTQIGAYFG 458  
Db 431 PLASYLGYTVNSATIPGDVLYIAGQRYNHTGQVVIYKMGEDGNINILQTLGGQIGSYFG 490  
Qy 459 ASLCSVDVDSNGSTDLVIGAPHY-----YEQTR-GQVSVCPPLRGQRA 502  
Db 491 SVLTIDIDKSDYTDLLVAGPMYMGTEKEEQKVYVYAVNTRFYQMSLEPIROTCCS 550  
Qy 503 RWOCDAVLYGEOGPMW-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLPFGTSG 561  
Db 551 SLKDNSTCKENKNEPCGARFGTAIAVKDLNVGDFNDVIGAPLEDDHAGAVYIYHG-SG 609  
Qy 562 SGISPSHSQRIAGSKSPRIQYFGQSLSGQDITMDGLVDLTVGAQGHVLLLRSPVLRV 621  
Db 610 KTIREAYAQRIPSGGDGKTLKFFQSIHGEMDLNGDGLTDVTIGGLGAALFWARDVAV 669  
Qy 622 KALMEENPREVARNVPCNDQVVKGEAG--EVRVCLHVO-KSTRDLREGQIQSVVTYD 678  
Db 670 KVTWNEFPKVNIOKQNCR---VEGKETVCINATMCFHVXKLSKESIIYADLQ----YR 722  
Qy 679 LALDSGRPHSRAVFNET-----KNSTRQTVLGLTQTCETLKLQLPNCI----- 723  
Db 723 VTLDSLQIISRSFSGTOERKQIQRNITVRESE-----CIRHSFYMLQK 765  
Qy 724 EDVPSPIVBLNLSLVGTSALFAGNLPVLAEADAQRLPTALPPEKNCNGNDNICQDDLSI 783  
Db 766 HFDODSVRVTLDNFMT-DPENG-----PVJODALPNSVHEHIIPFAXDCGKERCISDLTL 819  
Qy 784 TFSFMSLCLVVGQPRE-FNVTVTRNDGBDSYNTQVTFPPPLDLSYRKVSTLQONQSR 842  
Db 820 NVSTTEKSLIVKQCHDKFNVSUVTXKNGDSAYNTRTVQHSNPLIPSGIEBIOKQ--- 875  
Qy 843 SWRLACESASTEVSALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLL-LKAN 901  
Db 876 ----SCSEN-----QNITCRVGYFFLRAGETVTFKIIFOFNTHLSLSENAIHLSAT 922  
Qy 902 VTSENNMPTNKTBEQLELPVKYAV---YMWVTSHGVS-----KYLNFYASENTSR 950  
Db 923 SDSEEPLESLNDNEVNISSIPVKEVGLQFYSSASEHHISVAANETIPEFINST--EDIGN 980  
Qy 951 VMQHQYQVSNLQGRSLP---ISLVP-----LVPVRLNQTVIWD-----RP--- 987  
Db 981 EINVPYTIKRGHPMPPELOLSISFPNLADGYPVLYPIG-----WSSSDNVNCRPSL 1034  
Qy 988 -----QVTFPS-----ENLSSTCHTKERLPSHSDFLAELRKAPVVCISIAV 1027  
Db 1035 EDPFGINSKXOTISKSEVLKRGTIQCSSTC-----GVANITCSLLP 1077  
Qy 1028 CQRIQCDI-----PPFGIOBEF---NATLGNLSFDWYKTSNHNHLLIVSTABILFND 1078  
Db 1078 SDSLQVNVVSLLLWKPTF-IRAHFSSNLMLTLAGELK-----SENSSLTSSN----- 1123  
Qy 1079 VFTLLPGQGAVRQTEKVPPEVNPPLP-IVGSSVGLLALLIITAALYKLGPFKR 1136  
Db 1124 -----RKRELAIQISKDLGPRVPLWVILLSAFAGLLMLLILALWKIGPFKR 1172  
Qy 1137 QYKDMSE 1144  
Db 1173 PLKKMEK 1180

Search completed: June 7, 2004, 17:12:38  
Job time : 18.1123 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 40.0446 Seconds  
(without alignments)  
9084.693 Million cell updates/sec

Title: US-09-902-481b-1

Perfect score: 5956

Sequence: 1 MALRVLLTALTLCHGFNLD.....FKRQYKDMWSEGGPPGABPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_bacteriaph.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4459	74.9	1151	11 Q9J130	Q9j130 rattus norv
2	3976.5	66.8	1036	11 Q8CA73	Q8ca73 mus musculus
3	3824.5	64.2	920	6 Q28984	Q28984 sus scrofa
4	3514	59.0	1169	4 Q81VA6	Q81va6 homo sapien
5	3352.5	56.3	1169	11 Q9QXH4	Q9qxh4 mus musculus
6	3282.5	55.1	1161	11 Q9QYE7	Q9qye7 rattus norv
7	1545	25.9	1161	11 Q9WTV4	Q9wtv4 mus musculus
8	1534.5	25.8	1160	11 Q9R200	Q9r200 mus musculus
9	1419	23.8	1196	13 Q98TF1	Q98tf1 cyprinus ca
10	1368.5	23.0	1187	13 Q98TF0	Q98tf0 cyprinus ca
11	1361	22.9	1086	4 Q96HB1	Q96hb1 homo sapien
12	1277	21.4	927	6 Q8HZV0	Q8hzv0 bos taurus
13	1171.5	19.7	1167	11 Q88340	Q88340 rattus norv
14	1125.5	18.9	1167	11 Q88341	Q88341 rattus norv
15	1068.5	17.9	1171	13 Q82094	Q82094 gallus gall
16	1060.5	17.8	1168	11 Q71QC3	Q71qc3 mus musculus

17	1049	17.6	1038	11 Q8BS01	Q8bs01 mus musculus
18	1029.5	17.3	1160	6 Q8MKF4	Q8mkf4 felis silve
19	1020	17.1	895	11 Q9WUF8	Q9wuf8 mus sp. itg
20	865	14.5	348	4 Q8TES5	Q8tes5 homo sapien
21	852	14.3	1332	5 Q9BPQ8	Q9bpq8 halocynthia
22	816	13.7	205	11 Q63001	Q63001 rattus norv
23	753.5	12.7	780	13 Q66271	Q66271 xenopus lae
24	738	12.4	823	4 Q8W118	Q8w118 homo sapien
25	686.5	11.5	823	11 Q8CE84	Q8ce84 mus musculus
26	683	11.5	1032	11 Q61989	Q61989 mus musculus
27	651.5	10.9	257	11 Q8C270	Q8c270 mus musculus
28	643	10.8	1036	11 Q91YD5	Q91yd5 mus musculus
29	641.5	10.8	1033	6 Q95G33	Q95g33 bos taurus
30	637	10.7	1474	5 Q86G87	Q86g87 pseudoplusi
31	619.5	10.4	1041	5 Q9UB90	Q9ub90 lytechinus
32	614.5	10.3	1041	5 Q76378	Q76378 lytechinus
33	589	9.9	1034	13 Q98T77	Q98t77 gallus gall
34	579.5	9.7	1054	5 Q9U6S1	Q9u6s1 strongyloce
35	555.5	9.3	1053	11 Q80YP5	Q80yp5 mus musculus
36	550	9.2	1033	13 Q42598	Q42598 xenopus lae
37	546	9.2	1036	6 Q7YRP8	Q7yrp8 equus cabal
38	535.5	9.0	1016	13 Q91779	Q91779 xenopus lae
39	530	8.9	974	11 Q924W2	Q924w2 rattus norv
40	529	8.9	1073	11 Q8CC06	Q8cc06 mus musculus
41	529	8.9	1119	5 Q86G88	Q86g88 pseudoplusi
42	526	8.8	1047	6 Q9MZD6	Q9mzd6 bos taurus
43	525.5	8.8	1007	6 Q9GK48	Q9gk48 bos taurus
44	522.5	8.8	1132	11 Q80Z18	Q80zi8 mus musculus
45	514.5	8.6	1034	6 Q9TUM4	Q9tun4 oryctolagus

#### ALIGNMENTS

#### RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.  
AC Q9J130  
DF 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DE Integrin beta 2 alpha subunit.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fathallah D.M. St., Zerria K. Jr.;  
RT Cloning of the rat CD11b cDNA sequence."  
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF268593; AAF81280.1; -  
DR HSP; P11215; 1BHQ.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VMP\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS50234; VFWA; 1.  
SQ SEQUENCE 1151 AA; 126943 MW; 8P785695D4074CA5 CRC64;

Query Match 74.9%; Score 4459; DB 11; Length 1151;  
Best Local Similarity 73.5%; Pred. No. 2.9e-315;  
Matches 847; Conservative 144; Mismatches 160; Indels 2; Gaps 2;

QY 1 MALRVLLTALTALCHGFNLDNTENAMTFOENARGFGQSVVQLGSRVVGAPQEIIVANOR 60  
DB 1 MTLKALLATVLTCHGFNLDNTENAMTFOENARGFGQSVVQLGSRVVGAPQEIIVANOR 60  
QY 61 GSLYQCDYSTGSCBPIRQVPEAVNMGLSLAATTSPPOLLACGPTVHOTCSNTYVK 120  
DB 61 GALTQCDYSTGSCBPIRQVPEAVNMGLSLAATTSPPOLLACGPTVHOTCSNTYVK 120  
QY 121 GLCFLFGNLRQOQKPPALRGCPQEDSDIAPLIDGSGSIIPHDPRMKFVSTVMEQL 180  
DB 121 GLCYLFGNLRQOQKPPALRGCPQEDSDIAPLIDGSGSIIPHDPRMKFVSTVMEQL 180  
QY 181 KKSXTLFSLMQYSEBERTHFTFKFONNPNRSVLPKPTOLLGRTHATGIRKVVRELFN 240  
DB 181 QKSTLFSLMQYSEBERTHFTFKFONNPNRSVLPKPTOLLGRTHATGIRKVVRELFN 240  
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQEL 300  
DB 241 KINGARDNAAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQEL 300  
QY 301 NTIASKPRDHVQVNNFEALKTIONQIRKIPAIETGTGSSSPHEHMSQGFSAIT 360  
DB 301 DTIASKPAGHVFQVNNFEALKTIONQIRKIPAIETGTGSSSPHEHMSQGFSAIT 360  
QY 361 SNGPILLSVGSYDWNAGGVFLYTSKXSTFINMTVRVDSMDNDAYLGVAALIIILNRVQSLV 420  
DB 361 SNGPILLSVGSYDWNAGGVFLYTSKXSTFINMTVRVDSMDNDAYLGVAALIIILNRVQSLV 420  
QY 421 LGAPRYQHIGLVAFRONTGWESNANVKTGTGAYFGASLCSVDVDSNGSTDLVLIGAP 480  
DB 421 LGAPRYQHIGLVAFRONTGWESNANVKTGTGAYFGASLCSVDVDSNGSTDLVLIGAP 480  
QY 481 HYETQIRGGVSVCLPRGQARQWQCDVLYGEGQPMGRFGAALTIVLGVNVDGKLTDDVA 540  
DB 481 HYETQIRGGVSVCLPRGQARQWQCDVLYGEGQPMGRFGAALTIVLGVNVDGKLTDDVA 540  
QY 541 IGAPBEDNRGAVLHGTGSGSIPSHSRIAGSLSPRLQYFGQSLGQDLTMDGLV 600  
DB 541 IGAPBEDNRGAVLHGTGSGSIPSHSRIAGSLSPRLQYFGQSLGQDLTMDGLV 600  
QY 601 DLTVGAQGVHLLRSQVLRVKAIMEFNPREVARNVPECNDOVVKKEAGEVRCVLRVOK 660  
DB 601 DLTVGAQGVHLLRSQVLRVKAIMEFNPREVARNVPECNDOVVKKEAGEVRCVLRVOK 660  
QY 661 STDRLREGIQSVVYDLDLGRPHSRVAFNETKRNSTRQVLTGTCTBLKLQLP 720  
DB 661 STDRLREGIQSVVYDLDLGRPHSRVAFNETKRNSTRQVLTGTCTBLKLQLP 720  
QY 721 NCIEDPUSPIVLRNLSLGTPLSAPGNLRPLVLAQALFTALPFEKNCNDTICQDD 780  
DB 721 NCIEDPUSPIVLRNLSLGTPLSAPGNLRPLVLAQALFTALPFEKNCNDTICQDD 780  
QY 781 LSITFPMSLDCLVVGQPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRS 840  
DB 781 LSITFPMSLDCLVVGQPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRS 840  
QY 841 QRSWRLACESASTESVGLKSTSCSINRPIPENSEVTFNTITFDVDSKASLGKULLKA 900  
DB 841 QRSWRLACESASTESVGLKSTSCSINRPIPENSEVTFNTITFDVDSKASLGKULLKA 900  
QY 901 NVTSENNRPTNKTEFQLELPVXYAVVMVTVSHGVSTKVLNFTASNTSRVMOHOVQSN 960  
DB 901 NVTSENNRPTNKTEFQLELPVXYAVVMVTVSHGVSTKVLNFTASNTSRVMOHOVQSN 960  
QY 961 LGORSIPISLVLPVRLNQVITWDRPQVTFNSLSTCHTKERLPSHSDFLAELRKAPV 1020  
DB 961 LGORSIPISLVLPVRLNQVITWDRPQVTFNSLSTCHTKERLPSHSDFLAELRKAPV 1020  
QY 1021 VNCISVACORIQCDIPFGIOBEFNATLKNLSFDWYIKTSHNLLIVSTAILFNDSVP 1080  
DB 1021 VNCISVACORIQCDIPFGIOBEFNATLKNLSFDWYIKTSHNLLIVSTAILFNDSVP 1080  
QY 1081 TLLPGQGAFVRQSTETKVEFFEPVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYND 1140  
DB 1081 TLLPGQGAFVRQSTETKVEFFEPVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYND 1140

DB 1079 ALLPQGGTFCVCAQTEKVBPTVHNVPPLVGVSSGVLVLLALITAGLYKLGFFKRYND 1138  
QY 1141 MNSEGPPGABEQ 1153  
DB 1139 MNNEAGGDPGPQ 1151  
RESULT 2  
Q8CA73 PRELIMINARY; PRT; 1036 AA.  
ID Q8CA73;  
AC Q8CA73;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Integrin alpha M.  
GN ITGAM OR F730045J24RIK.  
OS Mus musculus (Mouse).  
OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CS7BL/6J; TISSUE=Spinal cord;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The PANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RL Nature 420:563-573 (2002).  
DR EMBL; AK039444; BAC30350.1; --  
DR PIR; PT0572; PT0572.  
DR PIR; PT0633; PT0633.  
DR PIR; PT0697; PT0697.  
DR MGD; MGI:96607; Itgam.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 1.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188E77 CRC64;  
Query Match 66.8%; Score 3976.5; DB 11; Length 1036;  
Best Local Similarity 66.8%; Pred. No. 3.3e-280;  
Matches 771; Conservative 120; Mismatches 144; Indels 119; Gaps 2;  
QY 1 MALRVLLTALTALCHGFNLDNTENAMTFOENARGFGQSVVQLGSRVVGAPQEIIVANOR 60  
DB 1 MTLKALLATVLTCHGFNLDNTENAMTFOENARGFGQSVVQLGSRVVGAPQEIIVANOR 60  
QY 61 GSLYQCDYSTGSCBPIRQVPEAVNMGLSLAATTSPPOLLACGPTVHOTCSNTYVK 120  
DB 61 GALTQCDYSTGSCBPIRQVPEAVNMGLSLAATTSPPOLLACGPTVHOTCSNTYVK 120  
QY 121 GLCFLFGNLRQOQKPPALRGCPQEDSDIAPLIDGSGSIIPHDPRMKFVSTVMEQL 180  
DB 121 GLCYLFGNLRQOQKPPALRGCPQEDSDIAPLIDGSGSIIPHDPRMKFVSTVMEQL 180  
QY 181 KKSXTLFSLMQYSEBERTHFTFKFONNPNRSVLPKPTOLLGRTHATGIRKVVRELFN 240  
DB 181 QKSTLFSLMQYSEBERTHFTFKFONNPNRSVLPKPTOLLGRTHATGIRKVVRELFN 240  
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQEL 300  
DB 241 KINGARDNAAKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQEL 300

QY 301 NTIASKPRDHPVQVNFALKTIONQRLBKIPALIGTQTGSSSSPEHMSQEGFSAIT 360  
 DB 301 DTIASKPAGHPVQVNFALNTIONQRLBKIPALIGTQTGSSSSPEHMSQEGFSAIT 360  
 QY 361 SNGPLSTVGSYDWAGVFLYTSKSTFTINMTRVDSMDAYLGAAYAILNRVQSIV 420  
 DB 361 SNGPLSGVSGPDWAGVFLYTSKSTFTINMTRVDSMDAYLGAAYAILNRVQSIV 420  
 QY 421 LGAPRYCHIGLVAMFRONTGWSNANVKTQIGAVFGASLCSVDVDSNGSDTLVLIGAP 480  
 DB 421 LGAPRYCHIGLVAMFRONTGWSNANVKTQIGAVFGASLCSVDVDSNGSDTLVLIGAP 480  
 QY 481 HYVEQTRGGQVSCVPLPRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVAVCDKLTIVA 540  
 DB 481 HYVEQTRGGQVSCVPLPRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVAVCDKLTIVA 540  
 QY 541 IGAPGEDNRGAVILFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMGHLV 600  
 DB 541 IGAPGEDNRGAVILFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMGHLV 600  
 QY 601 DLTVAQGHVLLRSLVLRVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLLVQK 660  
 DB 601 DLTVAQGHVLLRSLVLRVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLLVQK 660  
 QY 661 STRDLRQEQIOSVVTYDIALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLP 720  
 DB 661 STRDLRQEQIOSVVTYDIALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLP 720  
 QY 721 NCIEDPVSPIVLRNLSVGTLSAFGNLAPVLAEDAQALFALPFPEKNCNGNDNICQDD 780  
 DB 721 NCIEDPVSPIVLRNLSVGTLSAFGNLAPVLAEDAQALFALPFPEKNCNGNDNICQDD 780  
 QY 781 LSITFMSLDCLVGCGPREFNVTVVRNDGSDSYRTQVTFPFLDLSYKVKSTLQNRS 840  
 DB 781 LSITFMSLDCLVGCGPREFNVTVVRNDGSDSYRTQVTFPFLDLSYKVKSTLQNRS 840  
 QY 841 QRSWL-ACESASSTEVSGALSKTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLK 899  
 DB 841 QRSWL-ACESASSTEVSGALSKTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLK 899  
 QY 900 ANVTSNNMPTNKTKEFQLELPVKYAVMVVTSVSHGVSTVYLNFTASENTSRVMOQYQVS 959  
 DB 900 ANVTSNNMPTNKTKEFQLELPVKYAVMVVTSVSHGVSTVYLNFTASENTSRVMOQYQVS 959  
 QY 960 NLGQSLPISLVLVPLRNQTVIMDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKAP 1019  
 DB 960 NLGQSLPISLVLVPLRNQTVIMDRPQVTFSENLSSTCHTKERLPSHDSFLAELRKAP 1019  
 QY 1020 VVNCSTAVCQRIQCDIPFTGICEENATLKGNSLFDWYIKTSHNLLIYSTAELLFNDVS 1079  
 DB 1020 VVNCSTAVCQRIQCDIPFTGICEENATLKGNSLFDWYIKTSHNLLIYSTAELLFNDVS 1079  
 QY 1080 FTLLPQGGAFVRSQETKVEPEFVNPFLPIVGVSSVGGILLALITAAALYKLGFFKQYK 1139  
 DB 1080 FTLLPQGGAFVRSQETKVEPEFVNPFLPIVGVSSVGGILLALITAAALYKLGFFKQYK 1139  
 QY 1140 DMSGEGGPQAPQ 1153  
 DB 1023 DMSGEGGPQAPQ 1036

RESULT 3

Q28984

ID Q28984 PRELIMINARY; PRT; 920 AA.

AC Q28984; 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE CD11b (Fragment).

GN CD11b.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 NCBI\_TaxID=9823;

RN [1]  
 SEQUENCE FROM N.A.

RA Lee J.-K., Schook L.B., Rutherford M.S.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; U40072; AAB16869.1; -.  
 DR HSSP; F11215; 1BHO.

DR GO; GO:0008305; C:integrin complex; IEA.  
 DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.

DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
 DR InterPro; IPR000413; Integrin\_alpha.

DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 4.

DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS50234; VWF\_A; 1.

FT NON\_TER 920 920  
 FT NON\_TER 1 1

SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.2%; Score 3824.5; DB 6; Length 920;  
 Best Local Similarity 79.8%; Pred. No. 3.2e-269;

Matches 735; Conservative 77; Mismatches 108; Indels 1; Gaps 1;

QY 134 POKPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKGFSVTVMQLKKSKTLPALMOYS 193  
 DB 1 POKPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKGFSVTVMQLKKSKTLPALMOYS 193

QY 194 EEPRIHFTFKFQNNPNRSLVKPITQLLGHTRHATGIRKVVRELPNITNGARKNAKIL 253  
 DB 61 EEPRIHFTFKFQNNPNRSLVKPITQLLGHTRHATGIRKVVRELPNITNGARKNAKIL 253

QY 254 VIIDGKFGDPLGVEDVPEADREGVIRYVIGVDAPFSEKSRQELNTIASKPRDRHVF 313  
 DB 121 VIIDGKFGDPLGVEDVPEADREGVIRYVIGVDAPFSEKSRQELNTIASKPRDRHVF 313

QY 314 QVNFEALKTIONQRLBKIPALIGTQTGSSSSPEHMSQEGFSAITNGPLSTVGSYD 373  
 DB 181 QVNFEALKTIONQRLBKIPALIGTQTGSSSSPEHMSQEGFSAITNGPLSTVGSYD 373

QY 374 WAGGVFLYTSKSTFTINMTRVDSMDAYLGAAYAILNRVQSIVLGAAYAILNRVQSIV 433  
 DB 241 WAGGVFLYTSKSTFTINMTRVDSMDAYLGAAYAILNRVQSIVLGAAYAILNRVQSIV 433

QY 434 MFRONTGWSNANVKTQIGAVFGASLCSVDVDSNGSDTLVLIGAPHYTQTRGGQVSV 493  
 DB 301 MFRONTGWSNANVKTQIGAVFGASLCSVDVDSNGSDTLVLIGAPHYTQTRGGQVSV 493

QY 494 CPLPRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVAVCDKLTIVAIGAPGEDNRGAV 553  
 DB 361 CPLPRGQARWQCDVLYGEGQPGWGRFGAALTVLGDVAVCDKLTIVAIGAPGEDNRGAV 553

QY 554 YLFGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMGHLVLTVAQGHVFL 613  
 DB 420 YLFGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMGHLVLTVAQGHVFL 613

QY 614 RSQPVLRVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVOKSTRDLRQEQ 673  
 DB 480 RSQPVLRVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVOKSTRDLRQEQ 673

QY 674 VTYDIALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQPNCIDBPVSVILR 733  
 DB 540 VTYDIALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQPNCIDBPVSVILR 733

QY 734 LNPSLVCTPLSAPGNLAPVLAEDAQALFALPFPEKNCNGNDNICODDLSITFSPMSLDC 793  
 DB 600 LNPSLVCTPLSAPGNLAPVLAEDAQALFALPFPEKNCNGNDNICODDLSITFSPMSLDC 793

QY 794 VVGPREFNVTVVRNDGSDSYRTQVTFPFLDLSYKVKSTLQNRSQRSQRLACESASS 853  
 DB 794 VVGPREFNVTVVRNDGSDSYRTQVTFPFLDLSYKVKSTLQNRSQRSQRLACESASS 853



QY	DB	Sequence	QY	DB	Sequence
GN	DB	ITGAX.	547	DB	EEENRGAVYIFPHGASRODIAPSPSQRISASOIFSRIOYFGQSLSGGQDTRDGLVLA
OS	QY	Mus musculus (Mouse).	606	QY	AQGHVLLRSQVILRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVKQKSTRD
OC	DB	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	607	DB	SKGRVLLRTRPILRVSPVTHFTPAEISRSVFCQVAPQEQTLSDATVCLHIESPKTK
OC	DB	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	666	QY	LRBGIOQSVITYDALDSGRPHSPRAVNETKSTRQTVGLGTQTCETLKLQLPNCIED
NCBI_TaxID=10090;	DB	[1]	667	DB	L--GDLSTVTFDIALDHGRSLSTRAIFKTKTRALTRVKTGLKNTKCESVLLLPACVED
SEQUENCE FROM N.A.	QY	Huang X., Goreki K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,	726	QY	PVSPVLRNLNPSLVGTPLSAGNLRPLVLAEDAOQLFTALFPFPEKNCNDNI CODDLSITF
TSUCHIYA H.;	DB	RA	725	DB	SVTPTITLAPSLVGVPISSIQMLQPLAVDDCTYFASLPFEKNGADHI CODDLSVVF
"Isolation of Genes Selectively Expressed by Dendritic Cells.";	QY	RA	786	QY	SFMSLDCLVCGPREFNVTVTVANDGSDSYRTQVTPFPPLDLGSRKYSTLQ-----
Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.	DB	RA	785	DB	GFPDLKTLVGSDELAVNDVTGNDGSDSYGTTVTLFPVGLSFRVAVGQVFLRKEDQ
EMBL; AF211864; AAF23492.1; "	QY	RA	837	QY	--NORSORSKELACESASTEVSGALKSTSCSINHPIFPENSVETNITFDVDSKSLGN
HSSP; P11215; 1BHQ.	DB	RA	845	DB	QWQRGQSHLMCD--STPDRSQGLWSSTCSRRHVI FRGSGSTFLVTFDVSFKALGD
WGD; MGI:94609; Itgax.	QY	RA	895	QY	KLLKANVTSENMPRTNKTSFQLELPVKYAVYVTVSHGVSTKYLNFTASE-NTSRVMO
GO: 0008305; C:integrin complex; IEA.	DB	RA	903	DB	RLLLARVGSNNVPGTPKTTFQLELPVKYAVYVTVSHGVSTKYLNFTASE-NTSRVMO
GO: 0004895; F:cell adhesion receptor activity; IEA.	QY	RA	954	QY	HQYQVSNLQOESLPI SLAVFLAVPVRLNCTVLMWRPQVTFSENLSSTCHTKERLPSHSDLA
GO: 0007229; P:integrin-mediated signaling pathway; IEA.	DB	RA	963	DB	HRFQVNLGQDVEPSINFVWPIELKGEAVM-TVMVSHPONPVTQCYRNLKPTQFDLLT
InterPro; IPR000413; Integrin_alpha.	QY	RA	1014	QY	ELRKAIVVNCISIAVQRIODIPFFGQIEEFNATLKGNLISFDWYIKTSNHLHLLIVSTAEI
InterPro; IPR02035; VWF A.	DB	RA	1022	DB	HMKSQVPLDCSLADCLHLRCDIPSLGILDLYFLAGNLISFGWISQTLQKVKLLSEAEI
Pfam; PF00357; integrin_A; 1.	QY	RA	1074	QY	LFNDSVTLIPGQAFVRSQTEKTVPPFVNPPLPIVGVSSVGLLHLLALITLALYKLG
PRINTS; PR01185; INTEGRINA.	DB	RA	1082	DB	TFNTSVYSQLPGQAFVRSQTEKTVPPFVNPPLPIVGVSSVGLLHLLALITLALYKLG
PRINTS; PR00453; VWFADOMAIN.	QY	RA	1134	QY	FKROYKDWSE 1144
SMART; SM00191; Int alpha; 4.	DB	RA	1142	DB	FKROYKEMLE 1152
SMART; SM00327; VWA; 1.	QY	RA	RESULT 6		
PROSITE; PS00242; INTEGRIN_ALPHA; 1.	DB	RA	ID	Q90YB7	PRELIMINARY; PRT; 1161 AA.
PROSITE; PS00244; VWF A; 1.	QY	RA	AC	Q90YB7	
SEQUENCE 1169 AA; 129150 MW; C616412033219A6 CRC64;	DB	RA	DT	01-MAY-2000	(TrEMBLrel. 13, Created)
	QY	RA	DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)
	DB	RA	DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)
	QY	RA	DE	Alpha D integrin.	
	DB	RA	OS	Rattus norvegicus (Rat).	
	QY	RA	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	DB	RA	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
	QY	RA	OC	NCBI_TaxID=10116;	
	DB	RA	OC	[1]	
	QY	RA	RN	SEQUENCE FROM N.A.	
	DB	RA	RC	STRAIN-Sprague-Dawley;	
	QY	RA	RC	O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,	
	DB	RA	RA	Gallatin W.M.;	
	QY	RA	RA	"Cloning of rat alpha D, a novel beta 2 integrin.";	
	DB	RA	RL	Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.	
	QY	RA	RL	EMBL; AF021334; AAF2141.1; "	
	DB	RA	DR	HSSP; P11215; 1BHQ.	
	QY	RA	DR	GO: 0008305; C:integrin complex; IEA.	
	DB	RA	DR	GO: 0004895; F:cell adhesion receptor activity; IEA.	
	QY	RA	DR	GO: 0007229; P:integrin-mediated signaling pathway; IEA.	
	DB	RA	DR	InterPro; IPR000413; Integrin_alpha.	
	QY	RA	DR	InterPro; IPR02035; VWF A.	
	DB	RA	DR	Pfam; PF01839; FG-GAP; 3.	
	QY	RA	DR	Pfam; PF00357; integrin A; 1.	
	DB	RA	DR	EDNRGAVYIFPHGASRODIAPSPSQRISASOIFSRIOYFGQSLSGGQDTRDGLVLA	

```
DR Pfam: PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 22584913984A705E CRC64;

Query Match
Best Local Similarity 58.1%; Pred. No. 1.4e-229;
Matches 663; Conservative 158; Mismatches 306; Indels 15; Gaps 10;

QY 5 VLLIT--ALTLCGHNLDTEENATPFQENARGFGQSVVQLQGSVVVVGAPQEIIVAAANORG 62
DB VILLCGNVLASCHGSLNEDVEBEIVFREDAASFQIVVQFGSRLVVGAPLEAVANQTR 65
QY 63 LYQDYSTGSCBPIRLQVPEAVNMSLGLSLAATSPQQLACGPTVHTQTSYTVKGL 122
DB LYDCAPATGWCQPIVLSRPLEAVNMSLGLSLVATNNAQLACGPTAQRACVKNYAKGS 125
QY 123 CFLFGSLNRQKQKPFALRGCPQEDSDIAFLIDGSGIIPHDFRMRKPFVTVMBQLKK 182
DB CULLGSSL-QFQAVPASMEPCREQEMDIAFLIDGSGINQDFQKMDFFVXALMGEPAS 184
QY 183 SKTLPFLMOYSEFRTHFTKFKFQNNPNRSLVKPTQLLGRTHATGIRKVVRLFNIT 242
DB TSTLPFLMOYSNLTKHTFTFRFNILDPQSLVDPIVQLQGLTYTATGIRTVMBELFRSK 244
QY 243 NGARKNAFKILVITDGEKEDPLGVEDVIPEADREGVIRYVIGCDAPRSKRSQELNT 302
DB NGSRSAKILLVITDQKRPDLEYSVDIPADKAGIIRYAGVGDAPQSBTALKELNT 304
QY 303 IASKPPRDHVFOVNPPEALKTIQNLREKIFATEGTQTGSSSFHEMSQEGFSAAITSN 362
DB IGSAPPQDHVFKVGNFAALRSIQRLQEKI PALEGTQSRSSSFHEMSQEGFSALTSD 364
QY 363 GPLLSTVGSVDMAGGVFLTYSKEKSTFIMNTRVDSMDNDAYLGAAAIILNRVQSLVG 422
DB GPFVLGAVGSPSMGGAFLYPPNTRPTFINNSQENVDRSDSYLGYSTAVAFWGVHSLILG 424
QY 423 APRYQHGILVAMPFRONTGMWSESNVNVKGTQIGAYFGASLCSVDVDSNGSDTLVLICAPHY 482
DB APRHQTGKVIPTQBARHWRPKSEVVGRTQIGSYFGASLCSVDVDSGTDVLICAPHY 484
QY 483 YEYTRGGQSVVCLPRGQBARMCDAVLVEGQQPMGRFGAALTVLGDNVNGKLTDAVIG 542
DB YEYTRGGQSVVFPV-PGVGRWQCEATLHGEQHPMGRFGVALTVLGDVNGDNLDADVIG 543
QY 543 ARCEENRGAVLPHCTSGSGISPSQSORTAGSKLSPRLQYFGQSLGGQDLMGLVDL 602
DB APGEERSRGAVYIFHGASRLNMFSPSQRVGSQLSLRLQYFGQSLGGQDLMGLVDL 603
QY 603 TVGAQGHVLLLRQPVLRVKAIMFENPREVARNVFECNDQVVGKGEAGEVVRVCLHVQKST 662
DB AVGAQGHVLLLRSLPLLVKVELSRFAPMEVAKAVYQCWERTPTVLBAGEATVCLTVHKGS 663
QY 663 RDLRGQIQSVVTVYLDLSDGPHRAVFNETHNSTRTQTVGLHTQTCETILKQLPNC 722
DB PDLL--GNVQGSRYRLDALDPGLISRAIFDETCKNCTLTGRKTLGHDCHETVKLLPDC 721
QY 723 IEDVPSPILRLNFSLVGTPLSAFAGMLRPVLAEDAQRLFTALFPFPEKNCNDNICDDLS 782
DB VEDAVSPILRLNFSLVSDASAP-RNLHPVLAVGSDQHITASLPFPEKNCQELLCEGLG 780
QY 783 ITFSFMSLCLVVGGRPREFNVTVTVNDEGDSVRTQVTFPPFLDLISYRKVSTLONQSOR 842
DB ISFNFSGLQVLVVGSPELTVTVTVNDEGDSYGLVKEFYFAGLSYRRTVGTQ-QPHQY 839
QY 843 SWELACESASTVSGALKSTCSINHPIFENSEVENTITPDVDSKASLGNKLLKANY 902
DB FLRLACEAPEAQED--LRSSCSINHPIFREGAKTFTMTTVDVSYKATGLDELLLRAXA 897
```

```
QY 903 TSENMPTNKTERTOLELPYKYAVYVVTSHGVSTKVLNFTASENTSR-VVMOHOYQVSNL 961
DB SSENKPGDTNKTAFQLELPVAYITVITLISROEDSTNHFSSSGRRQEAHRYRVNVL 957
QY 962 QORSLPISLVELFVRLNQTVIMDRPQVTFSENLS--TCHTKERLPSHSDFLABELKAPV 1020
DB SPLKLAUVNFWPVLNGVAVMD---VTLSSPAQGVSCVSMKPPQNPFDLTQIRRSV 1014
QY 1021 VNCSTAVCORIQCDIPFPGIOEBENATLKGMLSPDWIKTSHNLLIIVSTAEIILFNDSVF 1080
DB LDCSIADCLHFRCDIPSDIQDELDFILRGNLSFGWVSQTLQEKVLLVSEAEITFDSVY 1074
QY 1081 TLLPQGAQFVRSQSTETKVEPEVENPFLPIVSGSVGGILLIALTAALYKLGFFKROYKO 1140
DB SOLPQGAFLRAQVETLEEVVYEPFLVAGSSVGGLLIALITVLYKLGFFKROYKE 1134
QY 1141 MW 1142
DB 1135 ML 1136

RESULT 7
Q9WTV4 PRELIMINARY; PRT; 1161 AA.
ID AC Q9WTV4;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha L.
GN ITGAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=DEA/23; TISSUE=Spleen;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065902; AAD25885.1; -.
DR HSSP; P20701; 11FA.
DR MGD; MGI:96606; Itgal.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007180; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VMP_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWA; 1.
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B2098431 CRC64;

Query Match
Best Local Similarity 25.9%; Score 1545; DB 11; Length 1161;
Matches 407; Conservative 219; Mismatches 462; Indels 108; Gaps 39;

QY 1 MALRY---LLLTALT---CHGFNLDTEENATPFQENARGFGQSVVQLQGSVVVVGAPQ 52
DB 1 MSFRIAGPRLLLLGLQLFAKAWSYNLDTRPTQSLAQAGRHFGYQVLQIEDG-VVVGAP 59
QY 53 EIVANQSGSIYQCDYSGSCPEIRLQVPEAVNMSLGLSLAATSPQQLACGPTVHTQ 112
DB 60 E---GDNTGGYHCTSEFCQPSLH--GSNHTSKYLGMTLATDAKGSLLACQGLRT 115
QY 113 CSENTYVVKGLCFGLSGNLRQKQPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKPF 172
```

Db 116 CDQNTYLSGLCYLPQSLGEPMLQNRPAVQECMKGVLDVLFVFDGSGSLDRKDFKILF 175  
Qy 173 VSTVMEQLKSKTLFSLMOYSEFRHFTFKF--QNNPNRSLVKETITOLLGRTHATGI 231  
Db 176 MKDVWRKLSNTSQFAVQSTDCRTEFTFLDYVQKNKPNPDVLLGVSQPMFLTTNFR 235  
Qy 232 RKVVRLEFNITNGARKNAFKILVITDGRKFGDPLGYEDVPEADREG-----VIRV 284  
Db 236 NYVAVHFKESGARPDAKVLVIITDG-----EASDKGNISAADITRYI 281  
Qy 285 IGVGDAPRSEKSRQELNTIASKPPDHVFOVANNPALKTIQNLREKIPAIETGTGSS 344  
Db 282 IGIKHFVSQKQTHIFASEPVEFVKILDTFELKDLFTDLQRIIVAIESTNRQDIT 341  
Qy 345 SPEHMSQEGFSAAITNSGFLSTVSGYDWAGVF--LYTSKEKSTFINMTRVDSMDNDAY 403  
Db 342 SFNMELSSGISADLSKGHAVGAKWAGGFLDLREDLQATFVGQELTSDVRGY 401  
Qy 404 LGYAAA--ILNRVQSLVIGAPRYOHIGLVAMFR--QNTGMWESNANVKGTQIGAVFGAS 460  
Db 402 LGYTVAMTSSRRLPILAGAPRYOHVQVLLFQAPAGGRWNQTKIETGTQISYFGE 461  
Qy 461 LCSVDVDSNGTDLVLIGAPHYETRGGOVSVCLPRGORARWQCDVLYGSGQGPWCR 520  
Db 462 LCSVDVDSNGTDLVLIGAPHYETRGGOVSVCLPRGORARWQCDVLYGSGQGPWCR 520  
Qy 521 FGAALTVLGVNDGKLTVAIGAPGEDNRGAVYLFHGTSGSGISPSHQZAGSKLSPR 580  
Db 519 FGAALTAITDINGDRUTDVAVAGPLEE--QGAVYIFNGKPG--GLSPQSORIQGAQVPG 575  
Qy 581 LQYFGSLSGQDLTWGLVLTGQGVHLLRSQVLEVRKAIMFENPREVARNPFCN 640  
Db 576 IRWFGSRHGVKDLGGDLADVVGEGRVVLSRPVVDVVTLSFSPEEIPVHEVCS 635  
Qy 641 DVVVKGEAG--EVRVGLHGVKSTDRRLREGQIOSVVTYDIALDSGRPHRAVFNENYKT 699  
Db 636 YSAREEQKHGVKLCACFRILKPLTPQ--FQGRLLANLSYTLQLOGRHMRSGFLPDGSHL 693  
Qy 700 RQOTVGLGLTQTCETLKLQPNIBDPVSHIVRLNFSV---GTPLSAFGN--LRPVLA 755  
Db 694 SGNSTITP--DKSLDFHFPFICIQDLISPIVNSLNFSLLEBEGTFRDQGRAMQPIRP 752  
Qy 756 DAQRLFTALPFPFKNCNNDNIQDDLSITPFSMLDCLVVGGE-----REFNVTVRN 809  
Db 753 SIHTV--TKEPFPKNCNNDNIQDDLSITPFSMLDCLVVGGE-----REFNVTVRN 809  
Qy 810 DGEDSVRTQVTFPDLSTVRKYSTIQNSQSRWRLACBSASTEVSGAL--KSTSCSIN 868  
Db 805 SGEDAYVWRLDLPFRGLSPKRVEMLO---PHSRMPVSCBEL--TEGSSLLTTLKCNYS 859  
Qy 869 HPFIPENSEVTNITFDVDSKASLGNKLLKANVTSEN--NMPRTNKTETQLELPVKYAV 927  
Db 860 SPIFKAGQVSLQVNFLLNWSWEDFVELNGTVHCHENENSSLOEDNSAATHIPVLYPN 919  
Qy 928 MVYTHGVSTKYLNFTASENTSRVMOHQYQV---SNLQSRSLFISLVFLVPLNQTIV 984  
Db 920 ILTKQENSTLXISFTPKGPKTOQVQHVQVRIQPSAYDHNM--TLEALVGP-----W 972  
Qy 985 DRPO--VTPSENLS---TCHTKE--RLPSHDFLAELKAPVNCIAVQRIQDIDF 1037  
Db 973 PHSEDPIYTWSDTDLVTRCHSEDLKRPSE---AEQPCPLPV-----QFRCPVIF 1021  
Qy 1038 FGIQEBFNATLKNLSPDWIKTSHNLLIVSTAEILFNDVSFTLLPFGQAFVRSQTEK 1097  
Db 1022 ---RRETLIQVTGVELSKIKAS--STLSLCSLSVSFNSSKHFLVGSKA--SEAQVLK 1076  
Qy 1098 VBEPFVNPLPLVGVSGVGLLILALITAYLYKLGFPKRYKDM--SEGGPQAE 1152  
Db 1077 VDLIHEKMLHVYVLSIGIGVILLFLITLALYKVGFFKRLKKEKMEADGVPNGSP 1132

RESULT 8  
Q9R200

ID Q9R200 PRELIMINARY; PRT; 1160 AA.  
AC Q9R200;  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)  
DE Integrin alpha L.  
GN ITGAL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1] TaxID=10090;  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spleen;  
RA Ma R.Z., Teuscher C.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF065901; AAD25884.1; --  
DR HSSP; P20701; 1LFA.  
DR MGD; MGI:96606; Itgal.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWFA; 1.  
DR PROSITE; PS00234; VWFA; 1.  
SQ SEQUENCE 1160 AA; 128127 MW; A3C531B139F1FAD CRC64;  
  
Query Match 25.8%; Score 1534.5; DB 11; Length 1160;  
Best Local Similarity 33.9%; Pred. No. 2,7e-102;  
Matches 405; Conservative 218; Mismatches 465; Indels 107; Gaps 38;  
  
Qy 1 MALRV-----LLITALT---CHGFNLDTENAMTPENA-RGFGQSVVLQGSFVVVGAQ 52  
Db 1 MSFRIAGPRLLLGLQLFAKWSYNLDRPTQSLAQAGRHFGYQVLIQEDG-VVVGAP 59  
Qy 53 EIVAAQNGSLYQCYSTGSCBPIRLQVEVAVNNSLGLSLAATSPOLLACGFTVHT 112  
Db 60 E---GNTGGLYCHTSSSEFCQPVSLH--GSNHTSKYLGLWTLATDAKGSLLACDPGLSRT 115  
Qy 113 CSNTYVYKGLCLFPGSNLRQPKPEALRGCPQSDSDIAFLIDSGSIIIPHDPRMKFP 172  
Db 116 CDQNTYLSGLCYLPQSLGEPMLQNRPAVQECMKGVLDVLFVFDGSGSLDRKDFKILF 175  
Qy 173 VSTVMEQLKSKTLFSLMOYSEFRHFTFKF--QNNPNRSLVKETITOLLGRTHATGI 231  
Db 176 MKDVWRKLSNTSQFAVQSTDCRTEFTFLDYVQKNKPNPDVLLGVSQPMFLTTNFR 235  
Qy 232 RKVVRLEFNITNGARKNAFKILVITDGRKFGDPLGYEDVPEADREG-----VIRV 284  
Db 236 NYVAVHFKESGARPDAKVLVIITDG-----EASDKGNISAADITRYI 281  
Qy 285 IGVGDAPRSEKSRQELNTIASKPPDHVFOVANNPALKTIQNLREKIPAIETGTGSS 344  
Db 282 IGIKHFVSQKQTHIFASEPVEFVKILDTFELKDLFTDLQRIIVAIESTNRQDIT 341  
Qy 345 SPEHMSQEGFSAAITNSGFLSTVSGYDWAGVF--LYTSKEKSTFINMTRVDSMDNDAY 403  
Db 342 SFNMELSSGISADLSKGHAVGAKWAGGFLDLREDLQATFVGQELTSDVRGY 401  
Qy 404 LGYAAA--ILNRVQSLVIGAPRYOHIGLVAMFR--QNTGMWESNANVKGTQIGAVFGAS 460  
Db 402 LGYTVAMTSSRRLPILAGAPRYOHVQVLLFQAPAGGRWNQTKIETGTQISYFGE 461  
Qy 461 LCSVDVDSNGTDLVLIGAPHYETRGGOVSVCLPRGORARWQCDVLYGSGQGPWCR 520

462 LCSVDLQDQGEALLIGAPLFFGQGRGVFTY---QROSLFEMVSELOQDGPVPLGR 518  
 521 FGAALTALVGVNGKLTVDVAIGABEDNRCANVYLPHGTSGSGISPSHSLQSLSPR 580  
 519 FGAALTALVGVNGKLTVDVAIGABEDNRCANVYLPHGTSGSGISPSHSLQSLSPR 575  
 581 LOYFGQSLGGQDUTMDGLVLTGAGCHVLLRSQPLRVKALMEFNPREVARNVFECH 640  
 576 IRWFGRIHGVKGLGGBELANVVGPRVVLSSRPVVDVTELSPSPREI PVHEVCS 635  
 641 DOVVKKEAG-EVRVCLHVOKSTRDLRLREGIOISVTVYDIALDSGRPHSRVAFNEKNST 699  
 636 YSAREEQHGVKACPRKPLTPQ--FOGRLANLSYTLQLDCHRMRSRGLFPDGSHEL 693  
 700 RROTVGLGTCCTBLKQLQPLNCIEDPVSPVLRNLPSLV---CTPLSARFN-LRPVLAE 755  
 694 SGNTSITP-DKSCLDLFFHPICLODLISPVNSLAFSLIEBGTFRDQGRAMQILRP 752  
 756 DAQSLFTALPFFKNCNNDNICODLSITFSFSLDCLVVGPP-----REFNVTVVRN 809  
 753 SIHTV-TKEIPFKNCGEDKKCEANLTLSPPARS-----GRLMSSASLAVETLSN 804  
 810 DGDSYRTQVTFPPPLDLYRKYSTLQNRQSRWRLACESASTEVSGAL-KSTSCIN 868  
 805 SGEDAYWRDLDPRLSPKAKVEMQ---PHSRMPVSCBEL--TEGSSLLTKLKNVS 859  
 869 HPIPPENSEVTNITFDVDSKASLGNKLLKANVTSEN-NMPRTNKTFFOLELPVYAVY 927  
 860 SPIPKAGQVSLQVMTLNSWEDFVNLGTVHCENSSQLQEDNSAATHIPVLPVN 919  
 928 MVTSHGVSTKYNLTASENTRVMOHYOVSNLGORSPLSLVFLVPLNQTWDRP 987  
 920 ILTKQENSTLYISFTPKGPKTQOHOHYQV---RIQPSAYDHNPET-LEALVGVRP 973  
 988 Q-----VTFSENLS--TCHTKE-RLPSSHSDFLAELRKAPVNCSTIAVQCIQCDIPFF 1038  
 974 HSEDLIITVNSVOTDPLVCHSEDLKPS-----SEABPCLPGV-QFRCPVIF- 1020  
 1039 GIOBEFNATLKNLSPWYIKTSHNLLIVSTABILFNDVFTLLPQOGAFVRSQTEKV 1098  
 1021 --RWEILLQVGTVELSKELKAS-STLSLCSLSVSNFSSKHPLHYGSKA-SEAQVLKV 1076  
 1099 EPFEVNPPLIVGVSSVGGILLALITAAALYKLGFFKQVQKMM-SECGPPGNAEP 1152  
 1077 DLHIEKEMLVVYVLSGIGGLVLLFLIPLALYKVGFFKELKEXKEADGGVPNGSP 1131  
 RESULT 9  
 Q98TF1 ID Q98TF1 PRELIMINARY; PROT; 1196 AA.  
 AC Q98TF1  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Coli-1.  
 GN C1A1.  
 OS Cyprinus carpio (Common carp).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Cyprinus.  
 OX NCBI\_TaxID=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Peritoneal exudate cells;  
 RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;  
 RT "Molecular cloning of a leukocyte integrin from the common carp.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR ENBL; AB048536; EMB39134.1; -.  
 DR HSSP; P20701; ILFA.  
 DR GO; GO:0008305; C:integrin complex; IEA.  
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.  
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.

DR GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001969; Asprotease\_A5.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; PG-GAP; 3.  
 DR Pfam; PF00357; integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PRO1185; INTEGRINA.  
 DR PRINTS; PRO0453; VWFADOMAIN.  
 DR SMART; SM00191; int\_alpha; 5.  
 DR SMART; SM00327; vwa; 1.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 DR PROSITE; PS0234; VWF\_A; 1.  
 SQ SEQUENCE 1196 AA; 132477 MW; 9369C807E7DCA53B CRC64;

Query Match 23.8%; Score 1419; DB 13; Length 1196;

Best Local Similarity 31.3%; Pred. No. 7.3e-94;

Matches 377; Conservative 228; Mismatches 455; Indels 144; Gaps 39;

3 LRVLLLTALTCLCHFNLDTENAMTFQENARG-FGQSVVQLQ-GSR--VVVGAPQBIIVAAAN 58

18 LGLILFMWASLGEAFNIDTEHPLRFGAPEDFFGYVQTEFGNRKQIIVGAPLE---GN 74

59 QRSGLYQCDYSTGSCPEIRKQVP-----VANVNSLGLSLAATSPOLLACQPTVHQTGS 114

75 STGEMYSCTADLQSCQ--RLQRPSESVRFGKSAVSAALTS-----CSPIYPHECD 126

115 ENTYYKGLCFLFGSNLRQOPQKFPFALRCQPEDSDIAPLIDGSGSIIPHDERRMKFVS 174

127 GNSYLVGVYQYSSSL-QAVSNFTAYQECSEKENVLFLPDGSSSMKTVDFEMKNFK 185

175 TYNEQLKQKGTFLSLMQISREPRHFTKPEPNNPNRELVKPITQLLGRTHATGIRKV 234

186 DIMKLSNSIIFAFAVQFSTDVRTVDFNDYQSGSAEKLME-THMSLTHTKAIDYI 244

235 VRELEN-ITNGARKNAFKLVITDGEKFGDPLGYED--VIPEADREGVIRVYVIGVDAP 291

245 LKNLNLMSLGSADSTAQALVITD-----GDPNDNDVNLKCKDEONILRIIVGVG--- 297

292 RSEKSRQELNTIASKPPRDHVFQVNFALKTIQNLREKIFAIETGTQTSSSSFEHMS 351

298 --KVDLIELTQLASEPKRNTFYIKDYGLKLLDNLQKKIYNIEGSDVAQCDRQKELS 355

352 QSGFSNAITSNGLPLSTGVSVDWAGVFLYTSKEKSTFINMTFRVDSDMN-DAYLVGAAL 410

356 QSGFSVYVQESVIVGSGVSNDRGALYEVTC-SGSDFAKETIIFAVNKGSGIMGTWT 414

411 ILERNVQSLVGLAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNG 470

415 GMRGVSLFLFGAPRAEHTGLVTLTKQNTVTVTSNNGEQIGSYFGASLSDLDVDSG 474

471 STDVLVIGAPHYEQTRGGQSVCLPRQGRARWCDVLYGEQ-----GQPKGFGAAL 525

475 DSDPLVGVAPLFYQSQ-----PRTEGRVYVYSLSEBQYFQKTLNVVSQSTTGFAASV 526

526 TVLGVGVNGDKLTVAIGAPGEDNRCANVYLPHGTSGSGISPSHS-QRIAGSKSLSPRLQVF 584

527 ASLADLNGDGLSDVAVGAPLE--NEGVMYVILGDRTHGINPELTQRISQSVLPCLQVF 584

585 GQSLSGQDGLTMDGLVLTGAGCHVLLRSQPLRVKALMEFNPREVARNVFECHDQV 644

585 GVSLTQGMNDNDNLTDVIGAGQGVILKAPVMSVAQLSPSPKBSISLNTFPCPGS-- 642

645 KGKEAGEVRVCLHVOKSTRDLRLREGIOISV--VTYDLALDSGRPHSRVAFNEKNSTR-- 700

643 NAFNANLTSCTVETRTSS---TGSLEKLVNSLVNVDVVRGMSRGFFDOSSVSSRTL 699

701 ROTQVGLTQTCETLKLQPLNCIEDPVSPVLRNLPSLVGTPLSPAGNLRPLVLAEDAKL 760

700 QQSVLLDSSGSCFNFSIFMLRCVADTVSPKIRNFS--QTEMLS-GNSVAVLDVHSRTE 756

761 FTALPPEFKNCNNDNICQDDLSITFSFMSLCLVVGPPREENVTVTVRNDGSDSVRTQV 820



```

Db 961 TTRVTLTPHNPMSQEVHHSFISSYHQ--LIIMCLNKLHFTSPBLSAVQTRTGRSLWVC- 1017
Qy 1030 RIQCDIPFGIOEFNATLKGNL-----SFDWYIKTSHNHLIIUSTAILF 1075
Db 1018 -----VSSISTGEIIFRSV--NLMAEAVLQNKVEYBSKISFY-EFRDRHVFNIS-AELNF 1068
Qy 1076 NDSVFTLLPG-QGAFVRSQSTETKVEPEVFNPLPIVGVSGVGLLLALITAAALYKLGFF 1134
Db 1069 NTSRYNSTGLKNPHSQTVEKVFEVIPPSRMLIVCTGAVGGFFFIILIIILLKCGFF 1128
Qy 1135 KROYKD 1140
Db 1129 KRNRPD 1134

RESULT 11
Q96HB1 PRELIMINARY; PRT; 1086 AA.
AC Q96HB1
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DB Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RX [1]
R2 SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; EC008777; AA008777.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; PG-GAP; 3_A.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02034; VWF_A; 1.
KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; F6FF2546B8C632F9 CRC64;

Query Match 22.9%; Score 1361; DB 4; Length 1086;
Best Local Similarity 32.0%; Pred. No. 1e-89;
Matches 381; Conservative 188; Mismatches 436; Indels 184; Gaps 38;

Qy 6 LLTALTLCGPF-----NLDENAMTFQ--ENARGFGQSVQLOQSRVVGVGAPQEI 54
Db 6 ITVMALLSGFFFFPASPNSYNDVRGARSFSPPRAGRHFGYVLQV-GNGVIVGAPGE- 63
Qy 55 VAANORSLYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPPQLACGTVHTQCS 114
Db 64 --GNSTGSLYQCSGSGCHCLPVTLR-CSNYSYKYLEMTLA--TDP----- 103
Qy 115 ENTVYKGLCLFGLNSLRQOPKQPPALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVS 174
Db 104 -----TDGS----- 107
Qy 175 TVMEQLKSKTFLSLMOYSBEFRIHFTFKEFQNNPNRSIAVKETITOLLGRTHATGIRKV 234
Db 108 -----ILFAAVQFSTSIKTEFDFSDIVRKQPDALKRKVXKMLLNTFGAINIV 157
Qy 235 VRELFNTNGARKNAFKILVITDGKFGDPLGYEVDIPEADREGVIRYVIGVDAPRSE 294
Db 150 ATEVFREELGARDPATKVLIIITDGE--ATDSGNIDAAXD-----IIRVIIGIKGHQTK 210

```

```

Qy 295 KSRQELNTIASKPPRDHYPOVNNFEALAKTIONLREKIFALIEGTOTGSSSEFEHMSORG 354
Db 211 ESQETLHKPASPASEFVKILDTFBLKLDLFTLQKIIYIEGTSKQDITFNMELUSSG 270
Qy 355 FSAATISNGPLSLTVSGSDVWAGVP--LYTSKEKSTFINNMTRVDSMDNDAYLAA--IIL 412
Db 271 ISADLSRGHAAVGVAGKDWAGGFLDLKADLQDDTFIGNELPLTPEVRAGLYGTVTWLPS 330
Qy 413 RNRVOSLVLCAPRYOHIGLVAMFR--QNTGMFESNANVKGTOIGAYFGASLCSVDSDNG 470
Db 331 ROKTSLASGAPRYOHMGVLLFQBPQGGHWSQVTHGTQIGSYFGELCGVDVDDQG 390
Qy 471 STDVLIGAPHYEQTRGGQSVSCPLERQCARWQCDV--LYGQGOQPMGFGAALTYL 528
Db 391 ETELLLIIGAPFYGRQGRVFYI-----QRQLGFEVSVLEQDGPYPLGRFGAITAL 445
Qy 529 GDVNGDKLTDVAIGAPGBEDNRGAVYLFHCTSGSGISPSHSORIASGSKLSPLOVFGQSL 588
Db 446 TIDNGDGLVDVAVGAPLEE--QGAVYIFNGRHG--GLSPQSORIEGTQVLSGIOMFGRSI 502
Qy 589 SGGQDLTMDGLVDLTVGAGHVLRLRSOPVLAVKAIMEFNPREVARNVPECDQV--KKG 647
Db 503 HGVKDLLEGDLADVAAGAESQMVILSSRPVDMVTLMSFPAEI PVHEVCCSVSTSNKMK 562
Qy 648 EAGEVRLVHVKOSTRDLRREGIOISVYTDIALDSGRPHSRVAVFNETKNTSTRQVILG 707
Db 563 EGVNITICFQI--KSLIPQF--QGRLVANLTYTLQLDGHRTREGLFPFGRHRLRNIAVT- 619
Qy 708 LTOTCETLKLQLPNCIEDPVSPVILRLNPSL---VDTPLSAPCN-----LRPLAEDAQ 758
Db 620 TSMSCDTSFHFPCVQODLISINVSINFLSWEBSGTFRDQAGKDIPILAPSLHSETW 679
Qy 759 RLFTALFPKKNCGNDNICODDLSITFSFMSLDCLVWGGPREFNTVTVNRNDEGDSYRTQ 818
Db 680 EI-----PFKNCGEDKCEANLRVSFSPARSALRLTAFASLSVELSLNLEEDAYWYQ 734
Qy 819 VTFFFPDLVYRKVSTLQNRORSORSLACES--ASSTEVSGALKSTSCSINHPIPPENS 876
Db 735 LDLPFPGLSFRKVMEL---KPHSQIPVSCCELPPESRLLSAL---SCNVSPIPKAGH 788
Qy 877 EYTFNITFDVDSKASLGKMLKANVTSENN----MPRTNKTFFOLELPVKVAVYVMTS 932
Db 789 SVALQMFNTLVNSSGSDVELHANVTGNEDSDLEONSATTI---IPILPINILIQD 845
Qy 933 HGVSTKYNFTASENTSRVMOHQOV---SNLQSSLP--ISLAVFLVPLVRLNQTWDRPQ 988
Db 846 QEDSTLYSFTPKGPKIHQVRKMYQVRIQPSIHDHNIPTLEAVVGVPQPPSGPITHQS 905
Qy 989 VTFSENLSSTCHK--RLIPSHSD--FLAELRKAFVNVNCSIAVCQRIQCDIPFGIOEFP 1044
Db 906 VQMBEPV--PCHVEDLERLPDAAEPCLPGALPRCPVW-----FRQEI 945
Qy 1045 NATLKGNSLFDWYIKTSHNHLIIUSTAILFNDVSFTLLPGQCAFVRSQETETKVEPFPV 1104
Db 946 LVQVITGLVELGVEIAS--SMFSLCSLSISFNSKHFHLYGSNLSL-AQVWKKVDVVEFK 1003
Qy 1105 NPLPIVGVSSVGGILLALLIATAALYKLGFPKQYKDMMSSEG-GPPGAEP 1152
Db 1004 QMLYLYVLSGIGLLALLIIFVLYKVGFEKELKKEWAGEGVNGIP 1052

RESULT 12
Q8HZV0 PRELIMINARY; PRT; 927 AA.
AC Q8HZV0
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lymphocyte function-associated antigen 1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

```

Ox NCBI_TaxID=9913;
Rn SEQUENCE FROM N.A.
Ra Thumbkat P., Kannan M.S., Maheswaran S.K.;
Rt "Sequence of the alpha subunit of bovine lymphocyte function-
Rl associated antigen 1.";
Rl Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
Dr EMBL; AF440778; AAN63636.1; -.
Dr PIR; A32039; A32039.
Dr GO; GO:0008305; C:integrin complex; IEA.
Dr GO; GO:0004895; F:cell adhesion receptor activity; IEA.
Dr GO; GO:0007160; P:cell-matrix adhesion; IEA.
Dr InterPro; IPR000413; Integrin_alpha.
Dr InterPro; IPR002035; VWF_A.
Dr Pfam; PF01839; FG-GAP; 3.
Dr PRINTS; PR01195; INTEGRINA.
Dr SMART; SM00191; Int.alpha.4.
Dr PROSITE; PS00242; INTEGRIN_ALPHA; 1.
Dr PROSITE; PS00234; VWFA; 1.
Dr NON_TER 1
FT TER 927
FT NON_TER 1
SQ SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;

Query Match 21.4%; Score 1277; DB 6; Length 927;
Best Local Similarity 34.8%; Pred. No. 1e-83;
Matches 340; Conservative 167; Mismatches 385; Indels 86; Gaps 28;

Qy 190 MQYSEPIHFTFEQNNPNRSLVKPIITOLLGRTHATGIRKVVRLPNIITGARNA 249
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 VQFSYFTEFTFDLYIQKDFDALLAGVKHLLTNTFGALNVAKVFPDVGARPD 63
Qy 250 FKILVITDGKFGDPLGVEDVIPADREGVIRVVGVDAPFRSKSQELNTIASKPPR 309
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
64 TKVLIITDGG---PPTNTLTMRPTSS---RSLGIGCKNFKTKESQALHQFASKVE 116
Qy 310 DRVFOVNNFEALKTIONLRKIPALEGTQTGSSSEFHEHMQEGFSAAITNSGPELLSTV 369
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 EFVKLIDTFEKLKDLTFEKLQKLYVIGTSGKDLTSFNNELSSGSDLSLSEGGHGVAG 176
Qy 370 GSYDWAGGVF-LYTSKSKSTFLNTRVDSMDNDAYLGAAA-IILNRVQSLVLGAPRYQ 427
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 GAKDWAGGFLLDKADLKSSTFVGNELTVESRAGYLVTVTLPSRGTSMLSLATGAPKYQ 236
Qy 428 HGLVAMFRO--NTGMWESNANVKTOIGAFGASLCSVDVDSNGSTDLVIGAPHYYEQ 485
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 HVGRLVFPQPRGGPWSQIOEIDGIGQISYFGGELCGVDVDRDGETELLIIAAPLYYGE 296
Qy 486 TRGGQSVSCPLPRGORARWQCDVLYGEOGQWGRFGAALTVLGVDNGDKLTDVAIGAPG 545
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
297 QRGGRVFIY---QKIQLEFQWYSELOGETGYPLGRFGAIAALTDINGDELTDVANGPL 353
Qy 546 EDNKGAVILFHGTSGSISPSHSORIASKLSPLRYFGQSLGQDLTDGGLVDLTVG 605
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
354 ES--QGVAVIFNGQQG--GLSPRPSORIBCTQMFSGIQWFGRSIHGKDLGGDGLADVAG 410
Qy 606 AQGHVLLARSOPVLAKVAMENPNREAVNPFECNDQVVKGEAG-EYRVCLHWK--ST 662
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
411 AEGQVIVLSSRVVDIIITSVSFPAEIPVHEVECSYSTNSQKKGUNLTVCFQWKSLLIST 470
Qy 663 RDLREGOIQSVVTVYDLALDGRPHSRVAFNPNKSTRQTVQLGLTCTETKLQLPNC 722
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
471 ----FQGHVLANVLTQLDGHRTSRGLFPGGKHLICNTAVTPV-KSCFVFWPHRPD 525
Qy 723 LEDPVSPIVLNPLS---VGTPLS--AFGNLRPVLAEDAQLFTALPFEKNGCNDNIC 777
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
526 IQDLISPIVNSLSYSLWBEEGFPDRALORDIPILKPSPHLETKETPFENKCEQKNC 585
Qy 778 QDDLITSPFMSLDCLVVGGPREFNVTIVRNDGDSYRTQVTFPFDLSTYKVTSTION 837
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
586 EADLKLAFSDMESKILRLTPSASLSVLRTAEDAAYVQVTLSPFGGLSFRKVEIL-- 643

```

```

838 QRSQSWRLACESASSSTEVSALKSTSCSINHPIPPENSEVTNITFDVDSKASLGKLL 897
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
544 -KPHSHVPVGCBLPEAAVHS-RAISCVSSSIFGDSWDIQVMENTLOKSGWGFIE 701
Qy 898 LKANVTS-----ENNMPRTNKTBEQLPVPKYAVYVMVTSHGVSSTKYLNFTASNTS 949
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
702 LQANVSCWNEEDSILLEDNATTS-----IPVMYPINVLTKQDENSTLYISPTPKSPRI 754
Qy 950 RVMOHQYQV---SNLQORSLSPLSLVPLVPRNLQTVI---WD---RPQVTFE-ENLSST 998
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
755 HHVKHYQVRIQPSNYDNMP-PLIALVVRVPRVHSEGLITKWSIQMEPPVNCSPRNLESP 813
Qy 999 CHTKERLPSSHDFLABELKAPVNVCSIAVCQRIQCDDIPFGIQQEENFNATLGNLSFDWYI 1058
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
814 SDEAE-----SCSPGT--EFCPCIDP--RQEILVQVNGMVLELRTGI 850
Qy 1059 KTSNHLIIVSTABILEPNDVSFTLLPGCGAPVRSOTETKVEPEVNPPLDILVSSVGG 1118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
851 KAS-SMILSCSSLAISFNKSHFHLHGRNASM-AQVVMKVDLVYEKMLYLYVLSGIGGL 908
Qy 1119 LLLALITAAALYKLGFFKR 1136
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
909 LLLFLIFIALYKVGFFKR 926

RESULT 13
O88340 PRELIMINARY; PRT; 1167 AA.
AC O88340;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin alpha E1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; Pubmed=9394838;
RA Brennan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
EL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020045; AAC23662.1; -.
DR HSP; P11215; IBBQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int.alpha.4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
SQ SEQUENCE 1167 AA; 128970 MW; D88A2C38ACDC2AAE CRC64;

Query Match 19.7%; Score 1171.5; DB 11; Length 1167;
Best Local Similarity 29.2%; Pred. No. 7.3e-76;
Matches 363; Conservative 213; Mismatches 464; Indels 205; Gaps 43;

Qy 5 VLLTALTALCHGNLDTENA--MTFQNRARGFGQSVVQLQGSRVVVGAPQEIIVANRGS 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
8 LLCMASLKPGCAFNMVDWAWVTALQPGAPVLSLHLDPN-----NOTCLIVARRSS 62
Qy 63 -----LYQCDYSNGSCPEIRLQVPEAVNMSLCLSLAATT--SPOLLAC-GPTVATC 113

```

Db	63	NRNTAALYRCALSI--SPDEIACQ--PVEHICMPKRGYGVTVLGNHNGVLVCIOVQARKFR	120
Qy	114	SENTYVKGCLFLPGSNLRQOPKPPALRG--	144
Db	121	SLNSLTGACSLTFLNLDLQAQAYFSDLEGFLDPGARVSDGYCRSGKGSTGEEKSARR	180
Qy	145	-----POEDSIAFLDGGSGSIIPEDPRMKFVSTVMEQL--KSKTLPFLMAYS	193
Db	181	RRVVEDEDEDETEIAVLVDGSGSEPSDFQAKNPFISTMMNFYEFKFCFPCNFPALVQVG	240
Qy	194	BEPRHIHTFKFQNNPNRSLVKPIITQLGRTHATGIRKVRVRELFNITNGARKNAFKIL	253
Db	241	AVIQTDFDLQESRDINASLAKVQSIQVKEVTKTASAMQHVLDNIPIPSGRGSKALKVM	300
Qy	254	VLTDEKFGDPLGYEDVPEADREGVIRVIGVGDAPRSEKSRQELNITASKPRDHPV	313
Db	301	VULTDGDIFGDLNLTIVNSPKQGVVPAIGVGDAPKNNYRELKLIADSPKEAHTF	360
Qy	314	QVNNFALKTIQOLREKIPAIEGTGTGSSSSFEHMSQEGFSAAITNSGP--LLSTVGSY	372
Db	361	KVTNYSALDGLLSKLOORIYHMEGT--VGDALQVQLAQTGFGSAILDKGVLGTGVGAF	417
Qy	373	DWAGGVPLY--TSKEKTFINMT--RVDS--DMNDAYLGVAAAILLRNRVQSLVIGAPRYQHI	429
Db	418	NWGGGALLYSTQNGRCFLNQTKADESRVTQVSYGLSVLAHLKAHGISYVAGAPHHKL	477
Qy	430	GLVAMFRQNTGMWESNA--NVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQT	486
Db	478	GAFFELKXEDR--EEDAFVRIISGEQMSYFGSVLCPVDINDGTTDFLVAAPFFHIRG	535
Qy	487	RGQVSVCLPRQARWQCDVLYGEQOPKRGFGAALTVLGDVNGDKLITVAGAP--	544
Db	536	EGRVVYQVPE--QDASFSLAHLTSGHPLGNTSRFGFNMAVGINQDKFTDVAIGAPLE	594
Qy	545	-----GEEDNRGAVYLPHTGSGSISPSHSORIASGLSPRLQVFGQSLSGGODLTMDGLV	600
Db	595	FGGAGDGASGVYIYNGHSG--GLYDSPSQIPRASSVAGSLHYFGNSVSGGLDFNGDGLA	653
Qy	601	DLTVGAGHVLRLRSQVPLRVKAIMFNPREVARNVFECNQDVVKGKAGEVVRVCLHVOK	660
Db	654	DIITVGRDRAVILSRPVDLTYSMTFTP-----DALPVPVIGKM--DYNLCFEVDS	703
Qy	661	S---TRDLREGIQSVVTVYDLALDSGRPHSRVAFNETKNSTRQTOVLGLTQC----	712
Db	704	SVVASEPGLREMFNTVDVV-----TKQRQLCEDSSGSCQCLURKWN	748
Qy	713	-----ETLKLQPLNCIEDPVSPVILRLNFSLVGTPLSAPGNLR-----PVLAED	756
Db	749	GGSPFLCEHFWLISTEEL---CEDCFSNITIKVYE-----FQTSGGRRDVPNPTL--D	797
Qy	757	AOELFTALP--PREKCGNDNTCODDLSITFSPMSLDCLVGCPREFNVTIVTRDGEDS	814
Db	798	HYKEPSAIFQLPFEKCKKRVPCIAEIQUTTN--ISQELVWGVKTEVTWNISLITNSGEDS	856
Qy	815	YRTOVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEVSGALKSTSCSINHPPEP	874
Db	857	YMINMALNYPRNLQPKKI-----QKPVSPDVQCDPKPV--ASVLVNNCKIGHPIIL--K	906
Qy	875	NSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTETFEQLELPVKYAVYVWVTSHG	934
Db	907	RSSVNVSVTWQLEESVFPNRTADITVITNSNEKSLARETR---SLOFPHAFIYLSR--	961
Qy	935	VSTKYLNFASENTSRVMQHOVQVSNLQORSILPSLVPLVPVRLNQTIVWDRPQVTFESN	994
Db	962	PSVMYWN--TSQSPSDHKPEFFNVHGENLFGAVFQIQVPIKIQDF-----QIVVRKN	1013
Qy	995	LSST-----CHTKERLPSHSDFLAEURKAPVNVCSIAVCORIQCDDIPPGIOEEFNATLK	1049
Db	1014	LTKTDHTCTEQSBACGSDPVQVHKWHSVVCAL-----TSNK	1053
Qy	1050	GNLSPDWYIKTSHNHLIVSTA-----ELLFNDVSFTLLPGQGVPRVQOTETKVEPF--	1101
Db	1054	ENVTVAAEISVGHTKQLLDVSELPIGLGELISFNKSLYEGLNAE-----NHRFKTITVFL	1107

Query Match 18.9%; Score 1125.5; DB 11; Length 1167;  
Best Local Similarity 27.8%; Pred. No. 1.6e-72;  
Matches 352; Conservative 214; Mismatches 456; Indels 243; Gaps 41;

Qy	5	VLLTALFLCHGNLDNENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIVAANQRGSLY	64
Db	6	LLCVASLPLGAFNVVDVNSW-----VTALQ-----FGAPSVL-----NSLIL	43
Qy	65	QCDYSTGSCPEIRLOVPVEAVNM--SLGLSLAATSPPOLLAGGPTVH-----	110
Db	44	QDPSNNQTC---LLVARRSSNRNSAVLYQCATSIIIDEGCQPVHEILMPKRGYQGVTL	99
Qy	111	-----OTCSENVYKGLCFLFGSNLRQCPQKPFPEALRG-----	143
Db	100	VRHNGVLVCIQVQSRKPRSLNSLTGACSLTNLDLQAOAYFSDLEGVLGASVNSG	159
Qy	144	-----CPOED-----SDIAFLDGGSGSIIPHDPRMKFVS	174
Db	160	DYGSKGGSTGEETKSARLRQAVBEDEBEACTEIAVLVDGSGSEPSDFQAKDPIS	219
Qy	175	TVMEQL--KSKTLPFLMAYSSEPRHIHTFKFQNNPNRSLVKPIITQLGRTHATGIR	232
Db	220	TMENFYEKCEPCFALVQGVVITQTEFDLLSDSDINASLAKVQSIQVKEVTKTASAMQ	279
Qy	233	KVRELFNITNGARKNAFKILVITDGEKFDPLGYEDVPEADREGVIRVIGVGDAPR	292
Db	280	HVLNDIIFPSRSGSKALKVMVLTGDIFEDPLNLTIVTSSSKMQGVVPAIGVGNAPF	339

RESULT 14

088341 PRELIMINARY; PRT; 1167 AA.  
AC 088341;  
DT 01-NOV-1998 (TREMELrel. 08, Created)  
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)  
DT 01-NOV-2003 (TREMELrel. 24, Last annotation update)  
DE Integrin alpha E2 (Fragment).  
DS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98056820; PubMed=9394838;  
EA Brenan M., Rees D.J.G.;  
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:  
RT tissue expression reveals phenotypic similarities between  
RT intraepithelial lymphocytes and dendritic cells in lymph.";  
RL Eur. J. Immunol. 27:3070-3079(1997).  
DR EMBL; AF020046; AAC23663.1; --  
DR HSSP; P11215; 1BHQ.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 4.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS50234; VWFA; 1.  
FT NON TER 1  
SQ SEQUENCE 1167 AA; 128593 MW; E3PED7E18B5CBEC CRC64;

```

QY 293 SEKSRQELNTIASKPPRHVFOVNFELKTIQNLREKIPAEIGTQTGSSSSPHEMSQ 352
DB 340 NNNTYRELKLASDPKAAHTFKVNYNSALDGLSKLQRIIHMEGT---VGDTLQQLAQ 396
QY 353 EGFSAITNSGP-LLSTGVSYNAGGVFLY-TSKSEKSTPINNTRVDSNDNA---YLGYA 407
DB 397 TGFSAQILDKGQVLLGTGAFNWSGALLYNTQNGRHFNLQT-AKEDFRAAQSYLQYS 455
QY 408 AAILRNVRQSLVLGAPRVOHILGLVAMPFRONTGMWESNANVSGTQIGAYFGASLCSVDVD 467
DB 456 VVALHKAHGVSVAGAPRHLKRGVAFELQKEDGETFMRIEKGWGSYFSGVSLCPVDIN 515
QY 468 SNGSTDLVLIGAPHYYEQTGROGVSVCPLEPRQORARWQCDVAVLYEGQOPWGRFGAALTU 527
DB 516 MDGITDFILVAAPPVHIREGKRVYIRV-HEQDAPFSLVYTLISGYPGLTSSRFQFMAA 574
QY 528 LGDVNGDKLTVAICAP-----GEEDNEGAVYLPHGTSGSGISPSHSORLAGSKLSPRL 581
DB 575 VGDINQDKFTDVAIGAPLEGFGAGDASGVSYYIYNGHSG-GHASPQOQIRASSVALGL 633
QY 582 QYFGOSLGGQDLTMDGLVDLTGAGQHVLLRLRSQPVLRVKAIMBPNPREVARNVFECDN 641
DB 634 YVFGMSVSGLDPSGDDLADITVGSQDVAVWLRSPVVDLTVSMFTPDALP----- 685
QY 642 QVVKGEAGEVRVCLHVQXS---TRDLREGQIQSVVTVYDLDLDSGRHRAVFNETGNS 698
DB 686 --MAFKDMVBLCFKVDSSAVPSEPLRGMSLNTFTVDV-----TKQK 728
QY 699 TRRQ-----TQVLGLTQTCETLKLQLENCIEDPVPVILRLNLSL 738
DB 729 QRLQCADRSCCCLMKWSSGSLCEHFGILSTEBEL-----CEDDCFSNITIKVSVEF 782
QY 739 VGTPLSAPGNLAPVLAARDQRLFTALP--PFRKNGDNICQDDLSITSPMSLCLLVG 796
DB 783 -QTSERRNHPNPL--DHYKEPSAIFQLPYEKCKXKVFCAEIQLTITAISQDD-LVVG 838
QY 797 GREENVTVTVNDGEDSVRTQVTPFPPLDLSYKSVSTLQONKORSQRLACESASSTEV 856
DB 839 IYKEVTMNLSTNSGSDSVWYNALNYPNLQFKI-----QKPLSPDIQCDPKPV-- 890
QY 857 SGALKSTCSINHPIPPENSEVTNFTPDVDVSKASIGNKLLKANVTSENMPRNKTEF 916
DB 891 -ASVLVNMCKIHPIL-KRSSVNVSVTQLESIFPNRTADITVITSNKSLSARETH- 947
QY 917 QLELPVXYAVMYVTVSHGVSSTKYNFTASGTSRVMOHOYOVSNLQORSLSPLSLVLPV 976
DB 948 --SLQPRHAFIAVLSR--FSVMYMN--TSQSSDHEKFEFFNVHGENHFAVQLQICVPI 1001
QY 977 RLNQTVINDRPQVTFSENLSST-CHTKERLPSHSDFLAELRKAPVNVCSIAVQRIQCDI 1035
DB 1002 TERDL-----QIIRVKHLTKTQANT-----ECTSQSEPTCGSDPVQNV---- 1039
QY 1036 PFGIOEENATL-----KGNLSFDWYIKTSHNHLII--VSTAEL-----FNDSVFTLL 1083
DB 1040 -----EHWHSVCAITSNKENTVAAEISMGHTKQLLRDISELQILGEISFNKSLYBGL 1093
QY 1084 PQGAFVRSQTEKVBEPF-----EVPNPLPLVGVSSVGLLALLALITAAALYKLGFKRQYK 1139
DB 1094 NAZ-----NHRKIIVIFLKEKPSLPLIIGSSIGGLLVLVVLIALLFKCGFFKRYK 1147
QY 1140 DMMSSE 1144
DB 1148 QLNLE 1152

```

## RESULT 15

O42094 PRELIMINARY; PRT; 1171 AA.

ID O42094

AC O42094;

DT 01-JAN-1998 (TEMBLrel. 05, Created)

DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)

DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)

DE ALPHAI Integrin.

```

OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus
OX NCBI_TaxID=9031;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=97476270; PubMed=9334246;
RA Obata H., Hayashi K., Nishida W., Momiyama T., Uchida A., Ochi T.,
RA Sobue K.;
RT "Sarcooth muscle cell phenotype-dependent transcriptional regulation of
RT the alphas integrin gene.";
RL J. Biol. Chem. 272:25643-26651(1997).
DR EMBL; AB000470; BAA23160.1; -
DR EMBL; AB000471; BAA23161.1; -
DR PIR; A55348; A55348.
DR HSSP; P17301; IAOX.
DR GO; GO:0003035; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWP_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;

Query Match 17.9%; Score 1068.5; DB 13; Length 1171;
Best Local Similarity 28.4%; Pred. No. 2.3e-68;
Matches 349; Conservative 215; Mismatches 521; Indels 145; Gaps 47;

QY 1 MALRVLLLTAL-TLCHGNLDTENAMTQENARG-FGOSVVQL---QGSRVVVGAPQIV 55
DB 1 MAASVWMLTALFHLTGASNVVDVKNMTSGPLEDMFGYTVQOYENEGKWLIGSLVQV 60
QY 56 AANQRGLSYQCDYSTGSCPE-IRLQVP-----VEAV--NMSLGLSAAATSPQLLAC 105
DB 61 PEKRTGDIYKCPVGRDSSQPCIKLNPDATSPVNVMEVKENMTLGTTL-VTNPKGGFLAC 119
QY 106 GPTVHTQTSNTYVYKGLCFLEGSNLRQOPQKPPALRCQPOEDSDIALIDGSGSIIPHD 165
DB 120 GPLYAYKGRGLHYTTGVCSNVSSSTPETVKAVAP-SVQEC-KTQDDIVIVLDGNSIYP-- 175
QY 166 FRRMKEFVSTWVEOLK--KSKTLPSLMQYSSEFRIHFTFKBFQNNPNRSLVKPTQLLG 223
DB 176 WESVTAFLNSILLRNWDIGPQQTQVGIQYGVTVVHEFYLVNTVSTTEEVMDAALRIQGG 235
QY 224 -RTHPTATGIRKVVRELNFNTNGARKNAFKIILVITDGEKFGDPLGYEDVIPEDREGVIR 282
DB 236 TQTMALGIDTAREAPTEAHGARRGVQKVMVIVTDGESH-DNYRLQEVIDKCEDENIQ 294
QY 283 YVIGVGDAF-----RSEKSRQELNTIASKPPRDHVQVNFPEALKTIQNLREKIFAIEG 337
DB 295 FAIALGSIYSGNLSTFKPVEEIKSIASKPTEKFFNVSDLEALVTIVELGCRIFALEA 354
QY 338 TQTGSSSSPHEHMSQGFSAAITNGPLLLSTVGSYDVAAGGVFLYTSKSEKSPFINNTRVDS 397
DB 355 ITDQQAASFEMEMSQGFSAHYSDQVWMLGAVGYDMNGTVMVWVDSDISPSNDFRDR 414
QY 398 DMND-----AVLGYAAAIIILNRVQSLVGLGAPRYCHIGLVAMFRQNTGMWESNANVKQ 452
DB 415 HSEKTEPLAALYGYTVNSALTPGGVLYTAGQPRNHTQVLIYKMGREKVLQRLKSEQ 474
QY 453 IGAYFGASLCSVDVDSNGSTDLVLIGAPHYY--EQTRGGQVSVCPFLPRG-----QRA 502

```

[illegible]

Search-completed: June 7, 2004, 17:16:09  
Job time : 48.0446 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 50.1568 seconds  
(without alignments)  
6495.175 Million cell updates/sec

Title: US-09-902-481B-1  
Perfect score: 5956  
Sequence: 1 MALRVLLTALTLCHGFNLD.....PFRQYKDMSEGGPGAEPPQ 1153

Scoring table: BLOSUMP2  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_25Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5956	100.0	1153	2	Aaw65090 Human Bet
2	5956	100.0	1153	3	Aab07360 Human CDI
3	5956	100.0	1153	5	Aau80252 Human Int
4	5956	100.0	1153	5	Abg61469 Human Bet
5	5956	100.0	1153	5	Aaol14428 Integrin
6	5956	100.0	1153	7	Add25615 Binding d
7	5946	99.8	1153	2	Aar04136 Alpha sub
8	3494	58.7	1163	2	Aar07120 p150.95 a
9	3480	58.4	1163	2	Aaw65091 Human Bet
10	3480	58.4	1163	3	Aab07361 Human CDI
11	3480	58.4	1163	5	Abg61470 Human Bet
12	3478	58.4	1163	6	Abu07406 Protein d
13	3455	58.0	1161	2	Aar78166 Human Bet
14	3455	58.0	1161	2	Aaw23049 Human Bet
15	3455	58.0	1161	2	Aaw57491 Human Bet
16	3455	58.0	1161	2	Aaw65089 Human Bet
17	3455	58.0	1161	2	Aaw72825 Human alp
18	3455	58.0	1161	2	Aaw73342 Human alp
19	3455	58.0	1161	3	Aab07359 Human alp
20	3455	58.0	1161	5	Abg61468 Human Bet
21	3439.5	57.7	1161	2	Aaw23064 Human Bet
22	3439.5	57.7	1161	2	Aaw65106 Human Bet
23	3439.5	57.7	1161	2	Aaw72837 Human alp
24	3439.5	57.7	1161	2	Aaw73343 Human alp
25	3439.5	57.7	1161	3	Aab07376 Human alp

26	3439.5	57.7	1161	5	ABG61485	Abg61485 Human Bet
27	3272.5	54.9	1161	2	AAR78169	Aar78169 Rat alpha
28	3270.5	54.9	1161	2	AAW23062	AAW23062 Rat beta
29	3270.5	54.9	1161	2	AAW60004	Aaw60004 Rat alpha
30	3270.5	54.9	1161	2	AAW72824	Aaw72824 Rat alpha
31	3270.5	54.9	1161	3	AAAB07374	Aab07374 Rat alpha
32	3270.5	54.9	1161	5	ABG61483	ABG61483 Rat Beta2
33	3264	54.8	1161	2	AAW23061	Aaw23061 Mouse bet
34	3264	54.8	1161	2	AAW60003	Aaw60003 Mouse alp
35	3264	54.8	1161	2	AAW65103	Aaw65103 Mouse bet
36	3264	54.8	1161	2	AAW72836	Aaw72836 Mouse alp
37	3264	54.8	1161	2	AAW73347	Aaw73347 Mouse alp
38	3264	54.8	1161	3	AAAB07373	Aab07373 Mouse alp
39	3263.5	54.8	1161	5	ABG61482	ABG61482 Mouse Bet
40	3263.5	54.8	1161	2	AAW65104	Aaw65104 Rat beta-
41	3263.5	54.8	1161	2	AAW73345	Aaw73345 Rat alpha
42	3260	54.7	1161	2	AAW78168	Aar78168 Mouse alp
43	3251.5	54.6	1151	2	AAW23059	Aaw23059 Rat beta
44	3251.5	54.6	1151	2	AAW60001	Aaw60001 Rat alpha
45	3251.5	54.6	1151	2	AAW65101	Aaw65101 Rat beta-

ALIGNMENTS

RESULT 1

AAW65090  
ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

XX DT 28-SEP-1998 (first entry)

XX DB Human Beta-integrin CD11b subunit protein.

XX KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
XX KW type-I diabetes; atherosclerosis; multiple sclerosis; aschma;  
XX KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;  
XX KW rheumatoid arthritis.

XX OS Homo sapiens.

XX FN US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-00485618.

XX PR 23-DEC-1993; 93US-00173497.

XX PR 05-AUG-1994; 94US-00286889.

XX PR 21-DEC-1994; 94US-00362652.

XX PA (ICOS-) ICOS CORP.

XX PI Van Der Vieren M, Gallatin WM;

XX DR WPI; 1998-206565/18.

XX PT Screening assay for modulators of integrin binding - using immobilised or

XX PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX PS Example 5; Fig 1A-D; 106pp; English.

XX CC This sequence represents a human beta-integrin CD11b subunit which is  
XX CC used to describe a method for identifying compounds that modulate the  
XX CC interaction of the beta-integrin alpha-d subunit with a binding partner  
XX CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
XX CC -d binding partner, one of which is immobilised and the other of which is  
XX CC labelled, in the presence of a test compound, and determining if the  
XX CC compound affects binding between the alpha-d polypeptide and alpha-d  
XX CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
XX CC comprising the cytoplasmic, transmembrane or extracellular domain of  
XX CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
CC and rheumatoid arthritis  
XX  
SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 2; Length 1153;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALTLCHGNLDENAMTFQENARGFGQSVVQLOGRVVGAPQBIVAANQR 60  
DB 1 MALRVLLTALTLCHGNLDENAMTFQENARGFGQSVVQLOGRVVGAPQBIVAANQR 60

QY 61 GSIYQCDYSGSCSEPTRLQVPEAVNMSLGLSLAATSPQLLACQPTVHQCSENTRYK 120  
DB 61 GSIYQCDYSGSCSEPTRLQVPEAVNMSLGLSLAATSPQLLACQPTVHQCSENTRYK 120

QY 121 GLCFLFGSNLRQOQPFPRALRCQPOEDSDIAFLIDSGSIIIPHDFFRMKEFVSTWEOQL 180  
DB 121 GLCFLFGSNLRQOQPFPRALRCQPOEDSDIAFLIDSGSIIIPHDFFRMKEFVSTWEOQL 180

QY 181 KKSRTLPSLMQYSEERPHHTFKEFQNNPNRSLVLPITQLLGRTHATGIRKVVRELPN 240  
DB 181 KKSRTLPSLMQYSEERPHHTFKEFQNNPNRSLVLPITQLLGRTHATGIRKVVRELPN 240

QY 241 ITNGARKNAFKILVITDGEKCDPLGYEDVPEADREGVIRVVGDAFSEKSRQEL 300  
DB 241 ITNGARKNAFKILVITDGEKCDPLGYEDVPEADREGVIRVVGDAFSEKSRQEL 300

QY 301 NTIASKPPRDHVPQVNNFEALKTIONLREKIPAEICTQTGSSSSFEHMSDEGFSAAIT 360  
DB 301 NTIASKPPRDHVPQVNNFEALKTIONLREKIPAEICTQTGSSSSFEHMSDEGFSAAIT 360

QY 361 SNGPLASTVGSYDMAGVPLTYSKEKSTFTNTRVDSMDNDVGLGYARAILLRNVQSLV 420  
DB 361 SNGPLASTVGSYDMAGVPLTYSKEKSTFTNTRVDSMDNDVGLGYARAILLRNVQSLV 420

QY 421 LGAPRYQHIGLVAMFRQNTGMWESSNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480  
DB 421 LGAPRYQHIGLVAMFRQNTGMWESSNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAP 480

QY 481 HYVEOTRGQVSVCPILPRGQARWQCDVLYGEGQPGWGRGNAALTVLGDVNGDKLTDVA 540  
DB 481 HYVEOTRGQVSVCPILPRGQARWQCDVLYGEGQPGWGRGNAALTVLGDVNGDKLTDVA 540

QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTMDGLV 600  
DB 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTMDGLV 600

QY 601 DLTVGAQCHVLLRSQVPLRVKIMFNPREVARNVPECDQVQVKGKAGEVRVCLHVQK 660  
DB 601 DLTVGAQCHVLLRSQVPLRVKIMFNPREVARNVPECDQVQVKGKAGEVRVCLHVQK 660

QY 661 STDRRLREGQIQSVVYTDALDSCRPHSRVAFNETKNSTRROTQVGLGTCTETLKLQLP 720  
DB 661 STDRRLREGQIQSVVYTDALDSCRPHSRVAFNETKNSTRROTQVGLGTCTETLKLQLP 720

QY 721 NCTEDVSPVILKNLPSLVGTPLSAFGNLRVLAEADAQRLFTALPFPERKNCNDNFCQDD 780  
DB 721 NCTEDVSPVILKNLPSLVGTPLSAFGNLRVLAEADAQRLFTALPFPERKNCNDNFCQDD 780

QY 781 LSTTFMSLDCLVWGSPREFNTVTVRNDGDSYRTQVTFPRPDLVSRYKYSTLQNGRS 840  
DB 781 LSTTFMSLDCLVWGSPREFNTVTVRNDGDSYRTQVTFPRPDLVSRYKYSTLQNGRS 840

QY 841 QRSWLACSSASTSVSGALKSTSCINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900  
DB 841 QRSWLACSSASTSVSGALKSTSCINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900

QY 901 NVTSENNMPRTNKTBEQLELPVKYAVYVMTVTSKVLTNFTASENTSRVMOHQYQVSN 960  
DB 901 NVTSENNMPRTNKTBEQLELPVKYAVYVMTVTSKVLTNFTASENTSRVMOHQYQVSN 960

QY 961 LQORSPLISLVFLVPLVRLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020  
DB 961 LQORSPLISLVFLVPLVRLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020

QY 1021 VNCISAVCORIOCDIPFGIOSEFNATLKGNSLFDWYIKTSHNHLIIYSTAEILLFNDVSF 1080  
DB 1021 VNCISAVCORIOCDIPFGIOSEFNATLKGNSLFDWYIKTSHNHLIIYSTAEILLFNDVSF 1080

QY 1081 TLLPQCGAFVRSQTEFKVEPEFVFNPLPLIIVGSSVGGLLMLALITAAALYKLGFFPRQVKD 1140  
DB 1081 TLLPQCGAFVRSQTEFKVEPEFVFNPLPLIIVGSSVGGLLMLALITAAALYKLGFFPRQVKD 1140

QY 1141 MMSDEGPPGAEPQ 1153  
DB 1141 MMSDEGPPGAEPQ 1153

RESULT 2  
AAB07360  
ID AAB07360 standard; protein; 1153 AA.  
XX  
AC AAB07360;  
XX  
DT 17-JAN-2001 (first entry)  
XX  
Human CD11b protein sequence.  
DE  
DB Human; macrophage infiltration inhibition; alpha\_d integrin;  
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;  
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;  
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;  
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;  
KW rheumatoid arthritis; central nervous system injury; CD11b.  
XX  
OS Homo sapiens.  
XX  
FN WO200029446-A1.  
XX  
PD 25-MAY-2000.  
XX  
PF 16-NOV-1999; 99WO-US027139.  
XX  
PR 16-NOV-1999; 98US-00193043.  
PR 08-JUL-1999; 99US-00350259.  
XX  
PA (ICOS-) ICOS CORP.  
XX  
PI Gallatin MW, Van Der Vieren M;  
XX  
DR WPI; 2000-387751/33.  
XX  
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit  
PT macrophage infiltration and reduce inflammation at central nervous system  
PT injury sites.  
XX  
PS Example 5; Fig 1; 270pp; English.  
XX  
CC Integrins are a class of membrane-associated molecules that participate  
CC in cellular adhesion. Integrins are made up of an alpha subunit and a  
CC beta subunit. One class of human integrins are restricted to expression  
CC in white blood cells and have a common beta2 subunit: the leukocyte  
CC integrins. Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins  
CC have an important role in immune and inflammatory responses. The present  
CC protein sequence is the human integrin alpha subunit CD11b. This sequence  
CC was used in an alignment to identify a novel beta2 integrin alpha  
CC subunit: alpha\_d (AAA60014 and AAB07359). The present sequence has  
CC approximately 50% identity to the protein sequence of alpha\_d. The  
CC Alpha\_d gene and protein may be useful in therapy for diseases linked to  
CC aberrant alpha\_d function e.g. Type I diabetes, atherosclerosis, multiple  
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory  
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency  
CC (LAD). In addition, anti-alpha\_d monoclonal antibodies may be used in the

CC inhibition of macrophage infiltration at the site of a central nervous  
CC system injury. The monoclonal antibodies can also be used to detect and  
CC diagnose Crohn's disease  
XX  
SQ Sequence 1153 AA;

```
Query Match      100.0%; Score 5956; DB 3; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCHGFNLDENAMTFQENARFGQSVVQLOGSRVVVCGAPQEIIVAAQR 60
DB 1 MALRVLLLTALTLCHGFNLDENAMTFQENARFGQSVVQLOGSRVVVCGAPQEIIVAAQR 60

QY 61 GSIYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATSPQLLACQPTVHQTCSENTYVK 120
DB 61 GSIYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATSPQLLACQPTVHQTCSENTYVK 120

QY 121 GICFLFGSNLRQOPQFPFALRCPOEDSDIAPLDGSGSIIPHDPFRMKFVFSTVMEQL 180
DB 121 GICFLFGSNLRQOPQFPFALRCPOEDSDIAPLDGSGSIIPHDPFRMKFVFSTVMEQL 180

QY 181 KKSCTLPSLMQYSEBRIHFTKPEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240
DB 181 KKSCTLPSLMQYSEBRIHFTKPEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240

QY 241 ITNGARKNAFKILVITDGEKCDPLGYEDVTPADREGVIRYVIGVDAFSEKSRQBL 300
DB 241 ITNGARKNAFKILVITDGEKCDPLGYEDVTPADREGVIRYVIGVDAFSEKSRQBL 300

QY 301 NTIASPPRDHVQVNFALQTKIQOLREKIPAEIGTGTGSSSPFEHMSQEGPSAALT 360
DB 301 NTIASPPRDHVQVNFALQTKIQOLREKIPAEIGTGTGSSSPFEHMSQEGPSAALT 360

QY 361 SNGPLLTSGVSDWAGVFLYTSKESFTINMTRVDSMDNDAYLGYAAAILLRNVQSILV 420
DB 361 SNGPLLTSGVSDWAGVFLYTSKESFTINMTRVDSMDNDAYLGYAAAILLRNVQSILV 420

QY 421 IGAAPRYQHILGVAMFQNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAP 480
DB 421 IGAAPRYQHILGVAMFQNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAP 480

QY 481 HYETQTRGQGVSVCPFRQARWQCDVLYGEGQGFQWGRFGAALTVLGDVNGDKLTDVA 540
DB 481 HYETQTRGQGVSVCPFRQARWQCDVLYGEGQGFQWGRFGAALTVLGDVNGDKLTDVA 540

QY 541 IGAPEEDNRGAVYLFHGTGSGISPSHSORAGSKLSPRLOYFGSLSGGQDLTMDGLV 600
DB 541 IGAPEEDNRGAVYLFHGTGSGISPSHSORAGSKLSPRLOYFGSLSGGQDLTMDGLV 600

QY 601 DLTVGAQGHVLLRSQPVLRVKALMEFNPREVARNVFECNDQVVRKEXAGEVVRVCLHWQK 660
DB 601 DLTVGAQGHVLLRSQPVLRVKALMEFNPREVARNVFECNDQVVRKEXAGEVVRVCLHWQK 660

QY 661 STRDLRREGIOQSVVYDIALDSGRPHSAVFNKTNSTRROTQVLGTOTCETLKLQLP 720
DB 661 STRDLRREGIOQSVVYDIALDSGRPHSAVFNKTNSTRROTQVLGTOTCETLKLQLP 720

QY 721 NCTEDPVSPIVLRNAPSLVGTPLSAFQNLRPVLAEDAQELFTALPPEKNCNDNICODD 780
DB 721 NCTEDPVSPIVLRNAPSLVGTPLSAFQNLRPVLAEDAQELFTALPPEKNCNDNICODD 780

QY 781 LSITPFSMLDCLVVGGPREFNVTVTRNDGSDSYNTQVTFPPDLDSYRKVSTLQNGRS 840
DB 781 LSITPFSMLDCLVVGGPREFNVTVTRNDGSDSYNTQVTFPPDLDSYRKVSTLQNGRS 840

QY 841 QRSWRLACSASTESVSGALKSTSCSINHPIPPENSEVTFNITPVDVSKASLGNKLLKA 900
DB 841 QRSWRLACSASTESVSGALKSTSCSINHPIPPENSEVTFNITPVDVSKASLGNKLLKA 900

QY 901 NVTSENMPRTNKTPEQLPLPKYAVYVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
DB 901 NVTSENMPRTNKTPEQLPLPKYAVYVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
```

```
QY 961 LQORSPLISLVPLVPLVRLNQTVIWDROPVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LQORSPLISLVPLVPLVRLNQTVIWDROPVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020

QY 1021 VNCIAVCORIQCDIPFFGIOEFNATLKGNSLSDWYIKTSENHLLIIVSTABILFNDSVF 1080
DB 1021 VNCIAVCORIQCDIPFFGIOEFNATLKGNSLSDWYIKTSENHLLIIVSTABILFNDSVF 1080

QY 1081 TLLPQGGAFVRSQETKVEPFEVNPPLPLIIVGSSVGGELLALITPAALYKLGFFKQYKD 1140
DB 1081 TLLPQGGAFVRSQETKVEPFEVNPPLPLIIVGSSVGGELLALITPAALYKLGFFKQYKD 1140

QY 1141 NMSEGGPPGABPQ 1153
DB 1141 NMSEGGPPGABPQ 1153

RESULT 3
AAU80252
ID AAU80252 standard; protein; 1153 AA.
XX
AC AAU80252;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human integrin 1 alpha-M subunit protein.
XX
KW Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 499..500
   /note= "Encoded by GGG CAG AGG"
XX
XX WO200218583-A2.
XX
XX 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US027227.
XX
XX 01-SEP-2000; 2000US-0229700P.
XX
XX (BLOO-) CENT BLOOD RES INC.
XX
XX Springer TA, Shimoaka M, Lu C;
XX
XX WPI; 2002-382964/41.
XX
XX N-PSDB; ASKS0046.
XX
XX Modified integrin-I or integrin I-like domain polypeptide useful as an
XX immunogen to produce antibodies specific to polypeptide, comprises a
XX disulfide bond such that polypeptide is stabilized in a desired
XX conformation.
XX
XX Disclosure; Page 109-112; 112pp; English.
XX
XX This invention relates to a modified integrin-I or integrin I-like domain
XX polypeptide comprising at least one disulfide bond so that the domain is
XX stabilised in a desired conformation. The polypeptide of the invention
XX may have antiinflammatory or immunosuppressive activities. The
XX polypeptides of the invention have an open conformation and are useful as
XX immunogens to produce antibodies that selectively bind to integrin I-
XX domain; and for identifying a modulator of integrin activity, or of
XX interaction of an integrin and a cognate ligand. The polypeptide of the
XX invention, or antibodies (preferably anti-LFA-1 antibody) is useful for
XX treating or preventing an integrin mediated disorder which is an
XX inflammatory or autoimmune disorder in a subject and for inhibiting the
XX binding of an integrin to a cognate ligand such as Crohn's disease,
```

CC nephritis; human immunodeficiency virus (HIV), myocardial infarction,  
 CC Sjoren's syndrome, rheumatoid arthritis, dermatitis. A therapeutic  
 CC composition comprising the peptide of the invention is useful for  
 CC treating an integrin mediated disorder in a subject. The polypeptides  
 CC and/or active or antigenic fragments are useful as reagents for diagnosis  
 CC of integrin-mediated disorders. The present sequence represents the human  
 CC integrin-1 alpha-M protein subunit used to generate the mutant  
 CC polypeptides of the invention  
 XX  
 SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 5; Length 1153;  
 Best Local Similarity 100.0%; Pred. NO. 0;  
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTCHGNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQIIVANQR 60  
 Db 1 MALRVLLLTALTCHGNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQIIVANQR 60

Qy 61 GSLYQCDYSTGSCPIPILOVPVEAVNMSGLSLAATSPDOLLACGPTVHQTSEITYVK 120  
 Db 61 GSLYQCDYSTGSCPIPILOVPVEAVNMSGLSLAATSPDOLLACGPTVHQTSEITYVK 120

Qy 121 GLCFPLGSLNRQPKPEALRGCPQSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQL 180  
 Db 121 GLCFPLGSLNRQPKPEALRGCPQSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQL 180

Qy 181 KSKTLFSLMOYSEEFRIHTFKEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELEN 240  
 Db 181 KSKTLFSLMOYSEEFRIHTFKEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELEN 240

Qy 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRVIVGVDAFRSEKSRQEL 300  
 Db 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRVIVGVDAFRSEKSRQEL 300

Qy 301 NTIASKPRDHVFOVNNPEALKTIONQREKIFAIEGTQTGSSSSFEHMSQSGPSAAIT 360  
 Db 301 NTIASKPRDHVFOVNNPEALKTIONQREKIFAIEGTQTGSSSSFEHMSQSGPSAAIT 360

Qy 361 SNGPLLSVTSYDWAGVGLVYTSKEKSTFINMTRVDSMDNDVLYGAAAILLRNVQSLV 420  
 Db 361 SNGPLLSVTSYDWAGVGLVYTSKEKSTFINMTRVDSMDNDVLYGAAAILLRNVQSLV 420

Qy 421 LGAPRYQHI GLVAMFRQNTGHWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480  
 Db 421 LGAPRYQHI GLVAMFRQNTGHWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480

Qy 481 HYETQTRGGQSVVCPPLRGQARWQCDVLYGQGPWGFGAALTVLGVNCGDKLTDVA 540  
 Db 481 HYETQTRGGQSVVCPPLRGQARWQCDVLYGQGPWGFGAALTVLGVNCGDKLTDVA 540

Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600  
 Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600

Qy 601 DLTVGAQGHVLLRSQVPLRVKAIMFNPNREAVRNPECDQVVKKEAGEVTVCLHVQK 660  
 Db 601 DLTVGAQGHVLLRSQVPLRVKAIMFNPNREAVRNPECDQVVKKEAGEVTVCLHVQK 660

Qy 661 STRDLREGQIQSVVYTDLALDSGRPHSRVAFNETKNSFRQTQVGLTQTCETLKLQLP 720  
 Db 661 STRDLREGQIQSVVYTDLALDSGRPHSRVAFNETKNSFRQTQVGLTQTCETLKLQLP 720

Qy 721 NCIEDPVSPIVLRNLSVGTPLSAFGLNPLVLAEDAQRLLFTALPFPEKNCQNDNTCQDD 780  
 Db 721 NCIEDPVSPIVLRNLSVGTPLSAFGLNPLVLAEDAQRLLFTALPFPEKNCQNDNTCQDD 780

Qy 781 LSITFSPMSLDCLVVGSPREFNTVTVRNDGDSYRQVTFPPDLVSRKVTSLNQKRS 840  
 Db 781 LSITFSPMSLDCLVVGSPREFNTVTVRNDGDSYRQVTFPPDLVSRKVTSLNQKRS 840

Qy 841 QRSWRLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900

Db 841 QRSWRLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900

Qy 901 NVTSENMPRTNKTETFOLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMQHOVQVSN 960

Db 901 NVTSENMPRTNKTETFOLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMQHOVQVSN 960

Qy 961 LQQRSLPISLVLVPLVRLNQTVIWDPRQVTPFSENISSCTHKERLPSPSHSDFLAELKAPV 1020

Db 961 LQQRSLPISLVLVPLVRLNQTVIWDPRQVTPFSENISSCTHKERLPSPSHSDFLAELKAPV 1020

Qy 1021 VNCSTAVCORIOCDIPFFGIQEEFNATLKNLSFDWYIKTSHNHLIIYSTAILFNDSVF 1080

Db 1021 VNCSTAVCORIOCDIPFFGIQEEFNATLKNLSFDWYIKTSHNHLIIYSTAILFNDSVF 1080

Qy 1081 TLLPQCGAFVRSQETETKVEPPEVNPPLPLIVGVSSVGGLLLALITAAALYKLGFFKQYKD 1140

Db 1081 TLLPQCGAFVRSQETETKVEPPEVNPPLPLIVGVSSVGGLLLALITAAALYKLGFFKQYKD 1140

Qy 1141 MMSBGGPPGABEQ 1153

Db 1141 MMSBGGPPGABEQ 1153

RESULT 4

ABG61469

ID ABG61469 standard; protein; 1153 AA.

XX AC ABG61469;

XX DT 27-AUG-2002 (first entry)

XX Human Beta2 integrin alphaCD11b subunit.

XX Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;

XX leukocyte adhesion deficiency; inflammatory response; diabetes;

XX multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;

XX inflammatory bowel disease; Crohn's disease; ulcerative colitis;

XX immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;

XX intracellular cell adhesion molecule; vascular cell adhesion molecule;

XX locomotor recovery; locomotor damage; locomotor impairment;

XX autonomic dysfunction; sensory dysfunction; spinal cord injury.

OS Homo sapiens.

XX WC200230980-A2.

XX PD 18-APR-2002.

XX PF 15-OCT-2001; 2001WO-US032059.

XX PR 13-OCT-2000; 2000US-00689307.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin WM, Van Der Vliet M;

XX DR WPI; 2002-463260/49.

XX Use of an anti-alpha-d monoclonal antibodies for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury.

XX Example 5; Page 191-194; 270pp; English.

XX The invention relates to promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following spinal cord injury by administering an anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a spinal cord injury victim. The method also involves the use of a ligand selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule, vascular cell adhesion molecule). The method is useful for promoting locomotor recovery, inhibiting locomotor damage, limiting locomotor impairment, or limiting autonomic and sensory dysfunction following

CC spinal cord injury. In particular, the spinal cord injury comprises  
 CC compression of the spinal cord. The antibodies are also useful for  
 CC reducing inflammation at the site of a central nervous system injury. The  
 CC specification also details the identification of Beta2 integrin alphas  
 CC cDNAs and proteins, for use in raising the antibodies. Beta2 integrins  
 CC are implicated in diseases such as LAD (leukocyte adhesion deficiency,  
 CC inflammatory response, diabetes, multiple sclerosis, arthritis, graft  
 CC atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative  
 CC colitis, immune complex alveolitis and leukaemia. The present sequence is  
 CC a Beta2 integrin alpha subunit sequence included for comparison with the  
 CC Beta2 integrin alpha protein sequences  
 XX  
 SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 5; Length 1153;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALFLCHGFNLDENATFQENARGFGQSVVQLGSRVVVGAPQEIIVANQR 60  
 DB 1 MALRVLLTALFLCHGFNLDENATFQENARGFGQSVVQLGSRVVVGAPQEIIVANQR 60  
 QY 61 GSLYQCDYSTGCEPTIRLOVPEAVNMSLGLSLAATSPQLLAGCTVHOTCSNTYVK 120  
 DB 61 GSLYQCDYSTGCEPTIRLOVPEAVNMSLGLSLAATSPQLLAGCTVHOTCSNTYVK 120  
 QY 121 GLCFLFGSLNRQQPQKPPALRGCCPOEDSDIAFLIDGSGSIIPHDFRMKEFVSTMQL 180  
 DB 121 GLCFLFGSLNRQQPQKPPALRGCCPOEDSDIAFLIDGSGSIIPHDFRMKEFVSTMQL 180  
 QY 181 KSKTLFSLMOYSEPRTHFTKPEFQNNPNPSLVKPTTOLLGRTHATGIRKVVRELFN 240  
 DB 181 KSKTLFSLMOYSEPRTHFTKPEFQNNPNPSLVKPTTOLLGRTHATGIRKVVRELFN 240  
 QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIFPADREGVIRYVIGVDAPSEKSRQEL 300  
 DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIFPADREGVIRYVIGVDAPSEKSRQEL 300  
 QY 301 NTIASKPPDRHVPQNNFPAKTIQNLREKIFAEGTGTGSSSPHEHMSOEGFSAIT 360  
 DB 301 NTIASKPPDRHVPQNNFPAKTIQNLREKIFAEGTGTGSSSPHEHMSOEGFSAIT 360  
 QY 361 SNGPLLSTVGSVDWAGGVFLYTSKEKSTFINTRVDSMDNDAYLGVAAILLRNVQSIV 420  
 DB 361 SNGPLLSTVGSVDWAGGVFLYTSKEKSTFINTRVDSMDNDAYLGVAAILLRNVQSIV 420  
 QY 421 LGAPRYQHIGLVAMFRQNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDLLVIGAP 480  
 DB 421 LGAPRYQHIGLVAMFRQNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDLLVIGAP 480  
 QY 481 HYYEOTRGQGVSCPLPRGORARWQCDVLYGEOGPWGRFGAALTVLGDVNGDKLTDVA 540  
 DB 481 HYYEOTRGQGVSCPLPRGORARWQCDVLYGEOGPWGRFGAALTVLGDVNGDKLTDVA 540  
 QY 541 IGAPGEDNRGAVYLPHTGSGSISPSHSORVAGSKLSPRLQYFGOSLSGGQDLTMDGLV 600  
 DB 541 IGAPGEDNRGAVYLPHTGSGSISPSHSORVAGSKLSPRLQYFGOSLSGGQDLTMDGLV 600  
 QY 601 DLTVGAQGHVLLRSQPLRVKAIMFPNREAVRVFECNDQVWVKEAGEVRVCLHVOK 660  
 DB 601 DLTVGAQGHVLLRSQPLRVKAIMFPNREAVRVFECNDQVWVKEAGEVRVCLHVOK 660  
 QY 661 STDRRLREQIQSVVYTDALDSGRPHSRVAFNETKNSTRRTQVGLTQTCETLKLQLP 720  
 DB 661 STDRRLREQIQSVVYTDALDSGRPHSRVAFNETKNSTRRTQVGLTQTCETLKLQLP 720  
 QY 721 NCIEDPVSIVLRLNLSLVTPLSAGNLRPVLAEDAQRFLTALFPFEKNCNDNICODD 780  
 DB 721 NCIEDPVSIVLRLNLSLVTPLSAGNLRPVLAEDAQRFLTALFPFEKNCNDNICODD 780  
 QY 781 LSITFFSMSLDCLVVGPREFNTVTVRVNDGDSYRTQVTFPPDLSTYRKVSTLQNRK 840  
 DB 781 LSITFFSMSLDCLVVGPREFNTVTVRVNDGDSYRTQVTFPPDLSTYRKVSTLQNRK 840

QY 841 QRSRLACSSASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLIKA 900  
 DB 841 QRSRLACSSASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLIKA 900  
 QY 901 NVTSENMPRTNKTETFOLELPVKYAVVMVTSHGVS TKYLNFTASENSTRVMQHOYQVSN 960  
 DB 901 NVTSENMPRTNKTETFOLELPVKYAVVMVTSHGVS TKYLNFTASENSTRVMQHOYQVSN 960  
 QY 961 LQQRSLPISLVFLVPVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020  
 DB 961 LQQRSLPISLVFLVPVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020  
 QY 1021 VNCISAVQCRIOCDIPFFGIQOEENATLKGNSLSPDWIKTSHNHLIVSTABILLPNSV 1080  
 DB 1021 VNCISAVQCRIOCDIPFFGIQOEENATLKGNSLSPDWIKTSHNHLIVSTABILLPNSV 1080  
 QY 1081 TLLPQCAFVRSQETETKVEPEVENPLPLIVGSSVGLLLALITAAALYKLGFFKQVYKD 1140  
 DB 1081 TLLPQCAFVRSQETETKVEPEVENPLPLIVGSSVGLLLALITAAALYKLGFFKQVYKD 1140  
 QY 1141 MMSEGGPPGAPQ 1153  
 DB 1141 MMSEGGPPGAPQ 1153  
 RESULT 5  
 AAO14428  
 ID AAO14428 standard; protein; 1153 AA.  
 XX AC AAO14428;  
 XX DT 03-MAY-2002 (first entry)  
 XX DE Integrin Mac-1 alpha subunit.  
 XX KW Mac-1; integrin alpha subunit; variant integrin inserted domain protein;  
 KW open conformation; integrin related inflammatory disorder;  
 KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;  
 KW reperfusion; hypovolemic shock; infarction; cerebral shock;  
 KW viral infection; cancer; gene therapy; vaccine;  
 KW bioactive agent screening.  
 XX OS Unidentified.  
 XX PN WO200204521-A2.  
 XX PD 17-JAN-2002.  
 XX PF 09-JUL-2001; 2001WO-US021805.  
 XX PR 07-JUL-2000; 2000US-0216600P.  
 XX PA (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
 XX PA (BLOO-) CENT BLOOD RES.  
 XX PI Springer T;  
 XX DR WPI; 2002-148167/19.  
 XX PT New integrin I domain protein having alteration in at least 2  
 PT noncontiguous regions and exits in an open conformation, useful for  
 PT treating, preventing or suppressing inflammatory or immunological  
 PT disorders.  
 XX PS Example 1; Fig 1F; 90pp; English.  
 XX CC The invention comprises structurally biased variant integrin inserted (I)  
 CC domain proteins, wherein the alterations to the protein occur in at least  
 CC two noncontiguous regions. Specifically the variant integrin I domain  
 CC proteins are structurally biased to exist in the open conformation,  
 CC thereby altering the binding ability of the protein. The invention also  
 CC comprises nucleic acids encoding the variant integrin I domain proteins.

CC The integrin I domain proteins and nucleic acids are useful for treating,  
CC preventing or suppressing integrin related inflammatory and immunological  
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain  
CC proteins and nucleic acids can also be used for treating: ischaemia/  
CC reperfusion (e.g. hypovolemic shock); infarction; cerebral ischemia/  
CC infection; and cancer. The variant integrin I domain nucleic acids and  
CC proteins may be used in gene therapy, as vaccines and to screen for  
CC bioactive agents. The present amino acid sequence represents the Mac-1  
CC alpha subunit of integrin  
XX  
SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 5; Length 1153;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALTLCCHGNLTENAMTFOENARGGSGWOLGSRVVVGAPQIVANOR 60  
DB 1 MALRVLLTALTLCCHGNLTENAMTFOENARGGSGWOLGSRVVVGAPQIVANOR 60  
QY 61 GSLVQCDYSTGSCPIRLQVPEAVNMSGLSLAATTSPQLLACGPTVHQTSENTRYK 120  
DB 61 GSLVQCDYSTGSCPIRLQVPEAVNMSGLSLAATTSPQLLACGPTVHQTSENTRYK 120  
QY 121 GLCFPLGSNLRQOKPEALRCGPQEDSDIAPLIDGSGSIIIPHDPRMKEFVSTWMEQL 180  
DB 121 GLCFPLGSNLRQOKPEALRCGPQEDSDIAPLIDGSGSIIIPHDPRMKEFVSTWMEQL 180  
QY 181 KSKXTLPSLMQYSEEFRIHTTFKEFQNNPNRSLVPEITQLLGRTHATGIRKVVRELPN 240  
DB 181 KSKXTLPSLMQYSEEFRIHTTFKEFQNNPNRSLVPEITQLLGRTHATGIRKVVRELPN 240  
QY 241 ITNGARKNAPKILVITDCKPGDPLGYEDVPEADREGVIRVYGVGDAPFRSEKROEL 300  
DB 241 ITNGARKNAPKILVITDCKPGDPLGYEDVPEADREGVIRVYGVGDAPFRSEKROEL 300  
QY 301 NTIASKPRDHVFOVNNFEALKTIQNLREKIPIAIEGTQTGSSSSPHEHNSQCFSAIT 360  
DB 301 NTIASKPRDHVFOVNNFEALKTIQNLREKIPIAIEGTQTGSSSSPHEHNSQCFSAIT 360  
QY 361 SNGPLLTSGSYDWAGGVFLYTSKEKSTFNTMRVDSMDNDAYLGVAALIIILNRVOSLV 420  
DB 361 SNGPLLTSGSYDWAGGVFLYTSKEKSTFNTMRVDSMDNDAYLGVAALIIILNRVOSLV 420  
QY 421 LGAPRYOHIGLVAMFRONTGWBNSANVKGITQIGAYFGASLCSDVDNSGSTDVLIGAP 480  
DB 421 LGAPRYOHIGLVAMFRONTGWBNSANVKGITQIGAYFGASLCSDVDNSGSTDVLIGAP 480  
QY 481 HYYETRGQGVSVCPPLRGQRARWQCDALVYEGQGPWGRFGAALTVLGVNMGDKLTDVA 540  
DB 481 HYYETRGQGVSVCPPLRGQRARWQCDALVYEGQGPWGRFGAALTVLGVNMGDKLTDVA 540  
QY 541 IGAPGEDNRCAVYLFHGTSGSGISPSHSORIASKLSPLQVFGQSLSGQDLTMDGLV 600  
DB 541 IGAPGEDNRCAVYLFHGTSGSGISPSHSORIASKLSPLQVFGQSLSGQDLTMDGLV 600  
QY 601 DLTVGAQGHVLLRSQPVLRKAIMENPREVANRVEPCNDQVVKGEAGEVRVCLHVQK 660  
DB 601 DLTVGAQGHVLLRSQPVLRKAIMENPREVANRVEPCNDQVVKGEAGEVRVCLHVQK 660  
QY 661 STRDLREGQOSVVTYDLALDSGRPHSRVAFNFKSTRTQTQVGLTCTCTKLQLP 720  
DB 661 STRDLREGQOSVVTYDLALDSGRPHSRVAFNFKSTRTQTQVGLTCTCTKLQLP 720  
QY 721 NCIEDPVSPIVRLNFSLVGTPPLSAFGNLRPLVAEDAQRLLFTALFPPEKNCQDNICQDD 780  
DB 721 NCIEDPVSPIVRLNFSLVGTPPLSAFGNLRPLVAEDAQRLLFTALFPPEKNCQDNICQDD 780  
QY 781 LSITFSFMSLDCLVVGGRPRENVTVVNDGEDSYRTQVTFPPPLDLSYKRVSTLQNRS 840  
DB 781 LSITFSFMSLDCLVVGGRPRENVTVVNDGEDSYRTQVTFPPPLDLSYKRVSTLQNRS 840  
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFTNITPDVDSKASLGKLLKA 900

DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFTNITPDVDSKASLGKLLKA 900  
QY 901 NVTSENNPRTNKTTFQLELPVKYAVYVWVTSKGYSTKVLNFTASENTRVMOHOVQSN 960  
DB 901 NVTSENNPRTNKTTFQLELPVKYAVYVWVTSKGYSTKVLNFTASENTRVMOHOVQSN 960  
QY 961 LQORSPLISLVLVPVRLNQTWIDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPV 1020  
DB 961 LQORSPLISLVLVPVRLNQTWIDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPV 1020  
QY 1021 VNCSTAVCQRIQCDIPPPGQIEEFNATLKGNLSPDWYIKTSHNHLIYSTABILFNDVSF 1080  
DB 1021 VNCSTAVCQRIQCDIPPPGQIEEFNATLKGNLSPDWYIKTSHNHLIYSTABILFNDVSF 1080  
QY 1081 TLLPQOGAFVRSQETKVEPEVPNPPLIYGVSSVGGLLALLAIIAALYKLGFFPKQYKD 1140  
DB 1081 TLLPQOGAFVRSQETKVEPEVPNPPLIYGVSSVGGLLALLAIIAALYKLGFFPKQYKD 1140  
QY 1141 MMSEGGPPGABEQ 1153  
DB 1141 MMSEGGPPGABEQ 1153

RESULT 6  
ADD25615  
ID ADD25615 standard; protein; 1153 AA.  
XX  
AC ADD25615;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Binding domain-immunoglobulin fusion protein-associated protein #85.  
XX  
KW Binding domain; immunoglobulin; fusion protein; cytostatic;  
KW antarthritic; immunosuppressive; antidiabetic; antithyroid;  
KW neuroprotective; hinge region; immunoglobulin heavy chain;  
KW CH2 constant region; CH3 constant region; IgG1;  
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
XX  
OS Unidentified.  
XX  
PN US2003118592-A1.  
XX  
PD 26-JUN-2003.  
XX  
XX 25-JUL-2002; 2002US-00207655.  
XX  
XX 17-JAN-2001; 2001US-0367358P.  
PR 17-JAN-2002; 2002US-00053530.  
PR 03-JUN-2002; 2002US-0385691P.  
XX  
PA (GENE-) GENE-CRAFT INC.  
XX  
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
XX  
XX WPI; 2003-801317/75.  
XX  
XX New binding domain-immunoglobulin fusion protein, useful for treating a  
PT subject having or suspected of having a malignant condition or a B-cell  
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
XX  
XX Disclosure; SEQ ID NO 176; 157pp; English.  
XX  
XX The invention relates to a binding domain-immunoglobulin fusion protein  
CC comprising a binding domain polypeptide that is fused to an  
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
CC CH2 constant region polypeptide that is fused to the hinge region  
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region  
CC polypeptide that is fused to the CH2 constant region polypeptide. The

CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
CC region polypeptide, derived from (a) having 3 or more cysteine residues;  
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
CC immunoglobulin hinge region polypeptide contains no more than one  
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
CC polypeptide, derived from (a) having 3 or more cysteine residues; where  
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
CC capable of at least one immunological activity comprising antibody  
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
CC binding domain polypeptide is capable of specifically binding to an  
CC antigen. Also included are an isolated polynucleotide encoding the  
CC binding domain-immunoglobulin fusion protein, a recombinant expression  
CC construct comprising the polynucleotide (operably linked to a promoter),  
CC a host cell transformed or transfected with a recombinant expression  
CC construct, producing the binding domain-immunoglobulin fusion protein, a  
CC pharmaceutical composition comprising the binding domain-immunoglobulin  
CC fusion protein or polynucleotide and a carrier, and treating a subject  
CC having or suspected of having a malignant condition or a B-cell disorder.  
CC The binding domain-immunoglobulin fusion protein is useful for treating a  
CC subject having or suspected of having a malignant condition or a B-cell  
CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
CC myasthenia gravis, Grave's disease, type 1 diabetes mellitus, multiple  
CC sclerosis or autoimmune disease. The present sequence is a binding domain  
CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
CC sequence data for this patent formed part of the printed specification  
CC and is also available in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocId=20030118592. The authors have not  
CC identified the sequences in the printed specification by their SEQ ID  
CC number therefore none of the sequences can be explicitly identified.

XX SQ Sequence 1153 AA;

Query Match 100.0%; Score 5956; DB 7; Length 1153;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALTLCHGFNLDENAMTFQENARGFQGVVQLGSRVVGAPQEIIVANQR 60  
DB 1 MALRVLLTALTLCHGFNLDENAMTFQENARGFQGVVQLGSRVVGAPQEIIVANQR 60  
QY 61 GSLYCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120  
DB 61 GSLYCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120  
QY 121 GLCFLFGSNLRQPKPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKGFVSTVMEQL 180  
DB 121 GLCFLFGSNLRQPKPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKGFVSTVMEQL 180  
QY 181 KSKTLFSLMOYSEFRTHFTFQNNPNPSLVKPIITOLLGRTHATGIRKVVRELPN 240  
DB 181 KSKTLFSLMOYSEFRTHFTFQNNPNPSLVKPIITOLLGRTHATGIRKVVRELPN 240  
QY 241 ITNGARKNAFKLVITDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQEL 300  
DB 241 ITNGARKNAFKLVITDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPFRSEKSRQEL 300  
QY 301 NTIASKPRDRHVFQNNPEALNTIQNLREKIPATEGCTQSSSSSFEHMSQGFSAIT 360  
DB 301 NTIASKPRDRHVFQNNPEALNTIQNLREKIPATEGCTQSSSSSFEHMSQGFSAIT 360  
QY 361 SNGPLLSVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLQYAAATILNRNVQSLV 420  
DB 361 SNGPLLSVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLQYAAATILNRNVQSLV 420  
QY 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAP 480  
DB 421 LGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAP 480

QY 481 HYETQTRGGQSVSCPLPRGQARWQCDAVLXGEQGPWGRFGAALTVLGVDVNGDKLTDVA 540  
DB 481 HYETQTRGGQSVSCPLPRGQARWQCDAVLXGEQGPWGRFGAALTVLGVDVNGDKLTDVA 540  
QY 541 IGAPGEEDNRGAVLFGHTSGSGISPSHSQRIAGSKLSPLOYEQCSLSGGQDLTMDGLV 600  
DB 541 IGAPGEEDNRGAVLFGHTSGSGISPSHSQRIAGSKLSPLOYEQCSLSGGQDLTMDGLV 600  
QY 601 DLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECDNVVVKGEAGEVRVCLHVQK 660  
DB 601 DLTVGAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFECDNVVVKGEAGEVRVCLHVQK 660  
QY 661 STRDLREGQIQSVVYDLDALDSGRPHSRAVFNTRKSTRTOVGLTQTCETLKLQLP 720  
DB 661 STRDLREGQIQSVVYDLDALDSGRPHSRAVFNTRKSTRTOVGLTQTCETLKLQLP 720  
QY 721 NCIEDPVSPIVRLNFSLVGTPLSAFGNLAPVLAEDAQRFLTALFPPEKNCNDNICQDD 780  
DB 721 NCIEDPVSPIVRLNFSLVGTPLSAFGNLAPVLAEDAQRFLTALFPPEKNCNDNICQDD 780  
QY 781 LSITFSFMSLDCLVVGGRBFNVTVVRNNGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840  
DB 781 LSITFSFMSLDCLVVGGRBFNVTVVRNNGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840  
QY 841 QRSWRLACESASSSTEVSALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKA 900  
DB 841 QRSWRLACESASSSTEVSALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKA 900  
QY 901 NVTSNNPRTNKTPEPOLELPVKVAVYVNTSHGVSTKYLNFTASENTRVWQHGYQVSN 960  
DB 901 NVTSNNPRTNKTPEPOLELPVKVAVYVNTSHGVSTKYLNFTASENTRVWQHGYQVSN 960  
QY 961 LGQSLPISLVFLVPVRLNQTIVMDRPOVTFSENLSSCTCHTKERLPSHSDFLAELRKAPV 1020  
DB 961 LGQSLPISLVFLVPVRLNQTIVMDRPOVTFSENLSSCTCHTKERLPSHSDFLAELRKAPV 1020  
QY 1021 VNCISAVCQRIQCDIPFFGIQEEFNATLKGNLSPDWYIKTSHNELLIVSTAEILFNDVSF 1080  
DB 1021 VNCISAVCQRIQCDIPFFGIQEEFNATLKGNLSPDWYIKTSHNELLIVSTAEILFNDVSF 1080  
QY 1081 TLLPGQAFVRSQTETKVEFEVNPPLIVGSSVGLLILALITAAALKLGPFRQYKD 1140  
DB 1081 TLLPGQAFVRSQTETKVEFEVNPPLIVGSSVGLLILALITAAALKLGPFRQYKD 1140  
QY 1141 MMSEGGPPGAEPO 1153  
DB 1141 MMSEGGPPGAEPO 1153  
RESULT 7  
AA04136  
ID AAR04136 standard; protein; 1153 AA.  
XX AAR04136;  
AC AAR04136;  
XX AAR04136;  
DT 25-MAR-2003 (revised)  
DT 07-SEP-1990 (first entry)  
XX Alpha subunit of Mac-1 leukocyte adhesion receptor.  
XX Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;  
KW non-specific defence system; integrin gene superfamily.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT Region 1..16  
FT /label= signal peptide  
FT Modified-site 86..88  
FT /label= putative N-glycosylation site  
FT Modified-site 240..242  
FT /label= putative N-glycosylation site  
FT Modified-site 391..393

FT	Modified-site	/label= putative N-glycosylation site	61	QY	GS	LYQCDYSTGSC	EP	IR	LQ	VP	VE	AV	NM	SL	GL	SL	LA	TT	SP	Q	L	AC	PT	V	H	Q	T	C	S	E	N	T	V	Y	K	120																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
FT	Modified-site	469..471	61	DB	GS	LYQCDYSTG	SC	EP	IR	LQ	VP	VE	AV	NM	SL	GL	SL	LA	TT	SP	Q	L	AC	PT	V	H	Q	T	C	S	E	N	T	V	Y	K	120																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
FT	Modified-site	/label= putative N-glycosylation site	121	QY	GL	CF	L	FG	SN	L	Q	Q	P	K	P	E	A	L	R	G	C	P	E	D	S	D	A	F	L	D	G	S	G	S	I	I	P	H	F	R	M	K	B	P	V	S	T	M	E	Q	L	180																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
FT	Modified-site	/label= putative N-glycosylation site	121	DB	GL	CF	L	FG	SN	L	Q	Q	P	K	P	E	A	L	R	G	C	P	E	D	S	D	A	F	L	D	G	S	G	S	I	I	P	H	F	R	M	K	B	P	V	S	T	M	E	Q	L	180																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
FT	Modified-site	/label= putative N-glycosylation site	181	QY	K	K	S	K	T	L	F	S	L	M	Q	Y	S	E	B	F	R	I	H	F	T	P	K	E	F	O	N	N	P	R	S	L	V	K	P	I	T	Q	L	L	G	R	T	H	T	A	T	G	I	R	K	V	R	E	L	P	N	240																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
FT	Modified-site	/label= putative N-glycosylation site	181	DB	K	K	S	K	T	L	F	S	L	M	Q	Y	S	E	B	F	R	I	H	F	T	P	K	E	F	O	N	N	P	R	S	L	V	K	P	I	T	Q	L	L	G	R	T	H	T	A	T	G	I	R	K	V	R	E	L	P	N	240																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
FT	Modified-site	/label= putative N-glycosylation site	241	QY	I	T	N	G	A	R	K	N	A	F	K	I	L	V	I	T	D	E	K	F	G	D	P	L	V	E	D	I	P	E	A	D	R	E	G	V	I	R	V	I	G	V	G	D	A	F	R	S	E	K	S	R	O	E	L	300																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
FT	Modified-site	/label= putative N-glycosylation site	241	DB	I	T	N	G	A	R	K	N	A	F	K	I	L	V	I	T	D	E	K	F	G	D	P	L	V	E	D	I	P	E	A	D	R	E	G	V	I	R	V	I	G	V	G	D	A	F	R	S	E	K	S	R	O	E	L	300																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
FT	Modified-site	/label= putative N-glycosylation site	301	QY	N	T	I	A	S	K	P	R	D	H	V	F	O	V	N	F	E	A	L	K	T	I	O	N	L	R	E	K	I	P	A	I	E	G	T	Q	T	S	S	S	S	F	H	E	M	S	O	E	G	F	S	A	A	I	T	360																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
FT	Modified-site	/label= putative N-glycosylation site	301	DB	N	T	I	A	S	K	P	R	D	H	V	F	O	V	N	F	E	A	L	K	T	I	O	N	L	R	E	K	I	P	A	I	E	G	T	Q	T	S	S	S	S	F	H	E	M	S	O	E	G	F	S	A	A	I	T	360																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
FT	Modified-site	/label= putative N-glycosylation site	361	QY	S	N	G	P	L	L	S	T	V	G	S	D	A	G	G	V	P	L	T	S	K	E	K	S	T	F	I	N	T	R	V	D	S	D	N	A	V	L	G	V	A	A	I	I	L	R	N	R	V	O	S	I	V	420																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
FT	Modified-site	/label= putative N-glycosylation site	361	DB	S	N	G	P	L	L	S	T	V	G	S	D	A	G	G	V	P	L	T	S	K	E	K	S	T	F	I	N	T	R	V	D	S	D	N	A	V	L	G	V	A	A	I	I	L	R	N	R	V	O	S	I	V	420																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
FT	Modified-site	/label= putative N-glycosylation site	421	QY	L	G	A	P	R	Y	Q	H	I	G	L	V	A	M	F	R	O	N	T	G	M	W	E	S	N	A	N	V	K	T	O	I	G	A	F	G	A	S	C	S	D	V	D	S	N	G	S	T	D	L	V	L	I	G	A	P	480																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
FT	Modified-site	/label= putative N-glycosylation site	421	DB	L	G	A	P	R	Y	Q	H	I	G	L	V	A	M	F	R	O	N	T	G	M	W	E	S	N	A	N	V	K	T	O	I	G	A	F	G	A	S	C	S	D	V	D	S	N	G	S	T	D	L	V	L	I	G	A	P	480																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
FT	Modified-site	/label= putative N-glycosylation site	481	QY	H	Y	E	Q	T	R	G	O	V	S	V	C	P	L	P	R	O	R	A	R	W	O	C	D	A	V	L	G	E	O	P	W	R	E	G	A	L	T	V	L	G	D	V	N	G	D	K	L	T	D	V	A	540																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
FT	Modified-site	/label= putative N-glycosylation site	481	DB	H	Y	E	Q	T	R	G	O	V	S	V	C	P	L	P	R	O	R	A	R	W	O	C	D	A	V	L	G	E	O	P	W	R	E	G	A	L	T	V	L	G	D	V	N	G	D	K	L	T	D	V	A	540																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
FT	Modified-site	/label= putative N-glycosylation site	541	QY	I	G	A	P	E	E	D	N	R	G	A	V	I	L	F	H	G	T	S	G	S	G	I	S	P	S	H	S	Q	R	I	A	G	S	K	L	S	P	R	L	Q	Y	F	Q	S	L	S	G	G	O	D	L	T	M	D	G	L	V	600																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
FT	Modified-site	/label= putative transmembrane region	541	DB	I	G	A	P	E	E	D	N	R	G	A	V	I	L	F	H	G	T	S	G	S	G	I	S	P	S	H	S	Q	R	I	A	G	S	K	L	S	P	R	L	Q	Y	F	Q	S	L	S	G	G	O	D	L	T	M	D	G	L	V	600																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
PN	EP364690-A.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										

New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating inflammation and viral infections, and in diagnosis.

Disclosure; Page ?; -pp; English.

Mac-1 alpha subunit is involved in the response to inflammation, i.e. recognition of and migration to sites of inflammation. It also attaches to cellular substrates as part of this function making it useful in visualising endothelial tissue. Mac-1 is a member of the Integrin Gene superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)

Sequence 1153 AA;

Query Match 99.8%; Score 5946; DB 2; Length 1153;  
Best Local Similarity 99.9%; Pred. NO. 0;  
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MALRVLLTALTLCGFNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVANQR 60  
1 MALRVLLTALTLCGFNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVANQR 60

Qy	1141	MMSEGGPPGAEPQ	1153
Dd	1141	MMSEGGPPGAEPQ	1153
 RESULT 8 AA07120 standard; protein; 1163 AA.			
ID	AA07120	standard; protein; 1163 AA.	
XX	AC	AA07120;	
XX	AC	AA07120;	
XX	DT	25-MAR-2003 (revised)	
XX	DT	05-FEB-1991 (first entry)	
DE	XX	p150.95 alpha subunit encoded by clone lambdaX47.	
KW	XX	p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia; rhinovirus.	
OS	XX	Synthetic.	
FH	Key	Location/Qualifiers	
FT	Region	1..19	
FT	/label=	signal peptide	
FT	Region	20..44	
FT	/label=	N-terminus	
FT	Modified-site	61..63	
FT	/label=	glycosylation site	
FT	Modified-site	89..91	
FT	/label=	glycosylation site	
FT	Modified-site	385..387	
FT	/label=	glycosylation site	
FT	Modified-site	392..394	
FT	/label=	glycosylation site	
FT	Modified-site	697..699	
FT	/label=	glycosylation site	
FT	Modified-site	735..737	
FT	/label=	glycosylation site	
FT	Modified-site	899..901	
FT	/label=	glycosylation site	
FT	Modified-site	904..906	
FT	/label=	glycosylation site	
FT	Modified-site	939..941	
FT	/label=	glycosylation site	
FT	Modified-site	1050..1052	
FT	/label=	glycosylation site	
FT	Domain	1108..1133	
FT	/label=	transmembrane	
XX	WO9010646-A.		
XX	PN		
XX	PD	20-SEP-1990.	
XX	PF	09-MAR-1990; 90WO-US001257.	
XX	PR	09-MAR-1990; 90WO-US001257.	
XX	PA	(DAND ) DANA FARBER CANCER INST INC.	
PI	Corbi AA,	Springer TA;	
DR	WPI; 1990-304985/40.		
DR	N-PSDB; AAQ06068.		
PT	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-unit of p150.95 cell surface adhesion receptor, opt. together with a beta chain of CD-18 family.		
PT	chain of CD-18 family.		
FS	Disclosure; Fig 3; 59pp; English.		
CC	Clone lambda X47 was isolated from a cDNA library constructed from total RNA extracted from phorbol myristate acetate stimulated HL-60 myelomonocytic cells. The library was screened with oligonucleotide probes based on tryptic peptide fragments of p150.95. The sequence can be attached to appropriate control elements and expressed in prokaryotic and eukaryotic cells. The protein can be used to treat or prevent rhinoviral infection because it interacts with ICAM-1 and inhibits cell-virus attachment. It can also be used as an anti-inflammatory agent. See also CC AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)		



Db 963 RDLFVSIINFPVPELVNQAEVWMDVSVFQNPFLSCSEKIAAPPDFAHIAQKNPVLDC 1022

QY 1024 STAVCORIQCDIPFGIOBENFATLKGMLSPDWYIKTSHNHLIIIVSTABILFNDSVFTLL 1083

Db 1023 STAGCLRFCDVPSVQVEELDTLKGMLSPGWVQIILQKXSVSVVAELIFDTSVYSQL 1082

QY 1084 PQGAFVRSQRTKTEPPEVPNPLPLIVGSSVGGLLILALITAAALYKLGPFKQYQYQMS 1143

Db 1083 PQCEAFMRAQTITVLEKYKVNPIPLIVGSSIGGLLLALITAVLYKVGPFKQYQYKME 1142

QY 1144 E 1144

Db 1143 E 1143

RESULT 10

ID AAB07361 standard; protein; 1163 AA.

XX AAB07361;

AC AAB07361;

XX 17-JAN-2001 (first entry)

DE Human CD11c protein sequence.

XX Human; macrophage infiltration inhibition; alpha d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; IAD; Type I diabetes;

KW atherosclerosis; multiple sclerosis; psoriasis; beta2 integrin;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11c.

XX Homo sapiens.

OS

XX WO200029446-A1.

PN

XX 25-MAY-2000.

PD

XX 16-NOV-1999; 99WO-US027139.

PF

XX 16-NOV-1998; 98US-00193043.

XX 08-JUL-1999; 99US-00350259.

PR

XX (ICOS-) ICOS CORP.

PA

XX Gallatin MM, Van Der Vieren M;

PI

XX WPI; 2000-397751/33.

DR

XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

PT macrophage infiltration and reduce inflammation at central nervous system

PT injury sites.

PF

XX Example 5; Fig 1; 270pp; English.

PS

XX Integrins are a class of membrane-associated molecules that participate

CC in cellular adhesion. Integrins are made up of an alpha subunit and a

CC beta subunit. One class of human integrins are restricted to expression

CC in white blood cells and have a common beta2 subunit: the leukocyte

CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins

CC have an important role in immune and inflammatory responses. The present

CC protein sequence is the human integrin alpha subunit C11c. This sequence

CC was used in an alignment to identify a novel beta2 integrin alpha

CC subunit: alpha d (AAB0014 and AAB07359). The present sequence has

CC approximately 56% identity to the protein sequence of alpha d. The

CC Alpha d gene and protein may be useful in therapy for diseases linked to

CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple

CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory

CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency

CC (IAD). In addition, anti-alpha d monoclonal antibodies may be used in the

CC inhibition of macrophage infiltration at the site of a central nervous

CC system injury. The monoclonal antibodies can also be used to detect and

CC diagnose Crohn's disease

XX SQ Sequence 1163 AA;

Query Match 58.4%; Score 3480; DB 3; Length 1163;

Best Local Similarity 61.1%; Pred. No. 6e-281;

Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;

QY 5 VLLTALTLCGFLDITENAMTFOENARGCQSVVLOGSRVYVGAQETVAANQORSGLY 64

Db 8 LLLFTALATSLGFLDITENAMTFOENARGCQSVVLOGSRVYVGAQETVAANQORSGLY 67

QY 65 QCDYSTGSCPIRLQVPEAVNMSLGLSLAATTPPQALLACGPTVHTQCSNTYKGLCF 124

Db 68 QCGYSTGACPIGLQVPEAVNMSLGLSLAATTPPQALLACGPTVHTQCSNTYKGLCF 127

QY 125 LFGNLRQOQKFPPEALRGCPQSDSDIAFLDGGSSIIIPDPRMKKFPVSTVMQLKSK 184

Db 128 LLGPT--QLTQRLVSRQECPEQDQIVFLDGGSSISRRNFATMMNPFRAVISQFORPS 185

QY 185 TLPISLMQYSEBFRIHTFKPEPPNNENPRSLVKPTITOLLGRTHATGKRVVRELFINITNG 244

Db 186 TQFSLMQSPKFTQHTFEEFRFTSNPLSLASVHQLGQFTYTATATQNVVHRLPHASVG 245

QY 245 ARKNAFKILVITDGEKFGDPIGYEDVIPEADREGVIRYIVGVGDAPSEKSRQELNTIA 304

Db 246 ARDAIKILVITDGEKFGDPIGYEDVIPEADREGVIRYIVGVGDAPSEKSRQELNTIA 305

QY 305 SKPRADHVQVNNPEALKTIONQLEKIFATEGTGSSSPSEHEMSOEGFSAATISNGP 364

Db 306 SKPSQEHIPKVEDFDALKDIONQLEKIFATEGTGSSSPSEHEMSOEGFSAATISNGP 365

QY 365 LLSTVGSYDWAAGVFLYTSKEKSTFINNTRVDSQNDAYLGVAALIIENRVQSVILGAP 424

Db 366 VLGAUGSPFTWGGAPLYPPNMSPTFINNQENQENQENQENQENQENQENQENQENQ 425

QY 425 RYQHILGLVAMFRONTGMEVSNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYE 484

Db 426 RYQHILGLVAMFRONTGMEVSNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYE 485

QY 485 QTRGGQVSVCPPLPGQARWOCDAVLYGEOQPGWRFGAALTVLGDVANGDKLTVAIGAP 544

Db 486 QTRGGQVSVCPPLPGQARWOCDAVLYGEOQPGWRFGAALTVLGDVANGDKLTVAIGAP 544

QY 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTWDGLVLT 604

Db 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTWDGLVLT 604

QY 605 GNOGHVLLLRQPVLRVKAIMEFNEPREVARNVFECNDQWKGKEAGEVVRVCLHVOKSTRD 664

Db 605 GARGQVLLLRQPVLRVKAIMEFNEPREVARNVFECNDQWKGKEAGEVVRVCLHVOKSTRD 664

QY 665 RLREGQIQSVVYDIALDSGRPHSAVFNENKNSRTRQTVLGLTQTCETLKLQPLNCIE 724

Db 665 RLREGQIQSVVYDIALDSGRPHSAVFNENKNSRTRQTVLGLTQTCETLKLQPLNCIE 724

QY 725 DVSPIVLRLNFSVGTPLSAPGNLRPVLAEDAQRLLTALPFRKNCNDNICODDLST 784

Db 725 DVSPIVLRLNFSVGTPLSAPGNLRPVLAEDAQRLLTALPFRKNCNDNICODDLST 784

QY 785 FSPFMSLDCVLWGGPREFNVTVTRNDGSDSYRTQVTFPPFLLDLSYRKVSTLQNSQSRW 844

Db 785 FSPFMSLDCVLWGGPREFNVTVTRNDGSDSYRTQVTFPPFLLDLSYRKVSTLQNSQSRW 844

QY 845 RLACASASTFVSGALKSTSCSIHPIPENSEVFNITFQVDSKASLGNKLLKANVTS 904

Db 845 RLACASASTFVSGALKSTSCSIHPIPENSEVFNITFQVDSKASLGNKLLKANVTS 902

QY 905 ENNMPTNKTEFQLELPVKYAVYVWVTSYKYNFTAS-ENTSRVMOHOYQVSNLQ 963

Db 903 ENNIPTSKTIFQLELPVKYAVYVWVTSYKYNFTAS-ENTSRVMOHOYQVSNLQ 962

QY 964 RSLPTSLVPLVVRNLQTVWDRPOVTFSENLSSTCTYKERLPSHSDFLAEARKKPVNC 1023

D5 963 RDLFVSNFVWPVNEQAVMVDVEVSHQPNQLRSCSEKIPASDPFLAHQKPNVLDLC 1022  
 QY 1024 SIACVQRTOCDIPFGIGQEEENATLKGNSLSDWYKTSNHHLLIVSTAEILENDVSTLL 1083  
 DB 1023 SIACVQRTOCDIPFGIGQEEENATLKGNSLSDWYKTSNHHLLIVSTAEILENDVSTLL 1082  
 QY 1084 PQGAFVRSOTETKVEPEVNPPLIVGSSVGLLALITAAIYKLGFFKQYKDMMS 1143  
 DB 1083 PQGAFVRSOTETKVEPEVNPPLIVGSSVGLLALITAAIYKLGFFKQYKDMMS 1142  
 QY 1144 E 1144  
 DB 1143 E 1143

RESULT 11  
 ABG61470  
 ID ABG61470 standard; protein; 1163 AA.  
 XX ABG61470;  
 AC  
 XX 27-AUG-2002 (first entry)  
 DT  
 DE Human Beta2 integrin alphaCD11c subunit.  
 XX  
 KW Beta2 integrin; alphaD subunit; CD11c subunit; CD11b subunit; LAD;  
 KW leukocyte adhesion deficiency; inflammatory response; diabetes;  
 KW multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;  
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;  
 KW immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;  
 KW intracellular cell adhesion molecule; vascular cell adhesion molecule;  
 KW locomotor recovery; locomotor damage; locomotor impairment;  
 KW autonomic dysfunction; sensory dysfunction; spinal cord injury.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200230980-A2.  
 XX  
 PD 18-APR-2002.  
 XX  
 PF 15-OCT-2001; 2001WO-US032059.  
 XX  
 PR 13-OCT-2000; 2000US-00688307.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 PI Gallatin WM, Van Der Vieren M;  
 XX  
 DR WPI; 2002-463260/49.  
 XX  
 PT Use of an anti-alpha-d monoclonal antibodies for promoting locomotor  
 PT recovery, inhibiting locomotor damage, limiting locomotor impairment, or  
 PT limiting autonomic and sensory dysfunction following spinal cord injury.  
 XX  
 PS Example 5; Page 194-198; 270pp; English.

The invention relates to promoting locomotor recovery, inhibiting  
 locomotor damage, limiting locomotor impairment, or limiting autonomic  
 and sensory dysfunction following spinal cord injury by administering an  
 anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a  
 spinal cord injury victim. The method also involves the use of a ligand  
 selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,  
 vascular cell adhesion molecule). The method is useful for promoting  
 locomotor recovery, inhibiting locomotor damage, limiting locomotor  
 impairment, or limiting autonomic and sensory dysfunction following  
 spinal cord injury. In particular, the spinal cord injury comprises  
 compression of the spinal cord. The antibodies are also useful for  
 reducing inflammation at the site of a central nervous system injury. The  
 specification also details the identification of Beta2 integrin alphaD  
 cDNAs and proteins, for use in raising the antibodies. Beta2 integrin  
 are implicated in diseases such as LAD (leukocyte adhesion deficiency,  
 inflammatory response, diabetes, multiple sclerosis, arthritis, graft  
 atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative

CC colitis, immune complex alveolitis and leukaemia. The present sequence is  
 CC a Beta2 integrin alpha subunit sequence included for comparison with the  
 CC Beta2 integrin alphaD protein sequences  
 XX  
 SQ Sequence 1163 AA;

Query Match 58.4%; Score 3480; DB 5; Length 1163;  
 Best Local Similarity 61.1%; Pred. No. 66-281;  
 Matches 697; Conservative 144; Mismatches 294; Indels 6; Gaps 4;  
 QY 5 VLLLTALTCHGFNLDTENAMTFQENARFGVQSVQLOQSRVVGVAPOEIVAAQNGSLY 64  
 DB 8 LLLFTALATSLGFLNLTDELTAFRVDSAGFGSVVQYANSWVVGAPQKILAAQIOGLY 67  
 QY 65 QCDYSTGSCBPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTQTCSENYKGLCP 124  
 DB 68 QCGYSTGACEPIGLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTQTCSENYKGLCP 127  
 QY 125 LFGSNLRQQQKFPFPEALRGCPQEDSDIAFLIDGSGIIIPDFFRMKEFTVSTVMEQKKSK 184  
 DB 128 LIGFT--QLTQLPVSRQCPQEQDIIVLIDGSGISSRNFAFMVFNRAVISQFQPS 185  
 QY 185 TLFLMYSSEBFRHFTPKBFQNNPNPSLVKPIITQLLGRTHATGIRKVRBELFNITNG 244  
 DB 186 TQPSLMQPSNKFTHTPEEFRTSNPLSLASVHQLQGTVTATAIQNVVHRLFHASYG 245  
 QY 245 ARKNAEKILAVITDCGKFGDPLGYEDVPEADREGVIRVYGVGDAFRSEKSRQELNTIA 304  
 DB 246 ARDAIKLITIVTDGKKGDSLDYKDVIPMAADAGIIRIYAGVGLAFQNRNSKELNDIA 305  
 QY 305 SKPPRDEHVFQNNFEALTKIQNLREKIPAIETGTQTGSSSSPEHEMSQGFSAATNSGP 364  
 DB 306 SKPSEHIFKVEDFDALDKIQNLREKIPAIETGTQTGSSSSPEHEMSQGFSAATNSGP 365  
 QY 365 LLSTGVSVDWAGGVLYTSKEKSTPINNTRVDSNDNDAYLGYAAAAILNRVQSLVIGAP 424  
 DB 366 VLGAAGSFTWGGAGFLYPPNMSPTFINMSQENVDNRDSYLGYSTELALWKGVQSLVIGAP 425  
 QY 425 RYOHIGLVAMFRONTGMSNSANVKGTOIGAFGASLCSVDVDSNGSTDLVLIGAPHYYE 484  
 DB 426 RYOHIGLVAMFRONTGMSNSANVKGTOIGAFGASLCSVDVDSNGSTDLVLIGAPHYYE 485  
 QY 485 QTRGQVSVCPILPRGQARWQCDALYQGGQPGWGFAGALTVLGDVNGDKLTDVAIGAP 544  
 DB 486 QTRGQVSVCPILPRGWR--RWCDALYQGGQPGWGFAGALTVLGDVNGDKLTDVAIGAP 544  
 QY 545 GEEDNRGAVTLFHTGTSGLSISPSHSQRIAGSKLSPLRLOYFGQSLSGGQDLTQDGLVDLAV 604  
 DB 545 GEEDNRGAVTLFHTGTSGLSISPSHSQRIAGSKLSPLRLOYFGQSLSGGQDLTQDGLVDLAV 604  
 QY 605 GAQGHVLLRSOPVLKVAIMEFPNPREVARNVPECDQVVKGEAGEVAVCLHVQKSTRD 664  
 DB 605 GARGQVLLRLTRPVLWVGVSQMFIPAIIPRSAPFCEBQVVSQTLVQSNICIIYIDKRSKN 664  
 QY 665 RLREGQIQSVVTVLALDSDGRPHSAVFNKSTNRRTQTVLGLTQTCETLKLQFNCIE 724  
 DB 665 RLREGQIQSVVTVLALDSDGRPHSAVFNKSTNRRTQTVLGLTQTCETLKLQFNCIE 724  
 QY 725 DPVSPVILRLNLSVLTPTLSAFGNLAPVLAEDAQRLFTALFFPEKXGNDNTCODDLST 784  
 DB 725 DSVIPILRLNLSVLTPTLSAFGNLAPVLAEDAQRLFTALFFPEKXGNDNTCODDLST 784  
 QY 785 FSPMSLDCLVGGSPRENTVTVNDGDSYRTVTFPPFLDLVSRKVSSTLQVRSQSW 844  
 DB 785 FSPMSLDCLVGGSPRENTVTVNDGDSYRTVTFPPFLDLVSRKVSSTLQVRSQSW 844  
 QY 845 RLACESASSTEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTS 904  
 DB 845 HLTC--CSAPVGSQGTWSTSCRINHPIFRGGAQITFLATFDVSPKAVGDLRLLIANVS 902  
 QY 905 ENNMPTNKTETQELPVPKYAVTMVVTSHGVSTKYNFTAS--ENTSRVVMQHOYQVNLSQ 963  
 DB 903 ENNMPTNKTETQELPVPKYAVTMVVTSHGVSTKYNFTAS--ENTSRVVMQHOYQVNLSQ 962

preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer permits the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer

Sequence 1163 AA;

Query Match 58.4%; Score 3478; DB 6; Length 1163;  
Best Local Similarity 61.0%; Pred. No. 8.8e-281;  
Matches 696; Conservative 136; Mismatches 303; Indels 6; Gaps 4;

QY 5 VLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAHQSGSLY 64  
DB 8 LLLFTALATSLGFWLDTTELTAFVDSAGFGDSVVQVANSVVVGAPQKITAANQTGGLY 67

QY 65 QDYSTGSCBPIRLQVPEAVNMSLGLSLAATSPQOLLACGPTVHOTCSENTVKGICLP 124  
DB 68 QCYSTGACBPIGLQVPEAVNMSLGLSLAATSPQOLLACGPTVHOTCSENTVKGICLP 127

QY 125 LFGSNLQQQKPPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKPFVSVTMQKSKK 184  
DB 128 LIGFT-QLTQRLPVSRQCEPRQDIVFLIDGSGSISSRNFAFMNFRVAVISQFQPS 185

QY 185 TLFSLMQVSEBPIRHTFKFQNNPNRSLVKPITQLGRTHATGKRVVRELPNTNG 244  
DB 186 TQSLMQVSEBPIRHTFKFQNNPNRSLVKPITQLGRTHATGKRVVRELPNTNG 245

QY 245 ARKNAKILVITDGEKFGPLGYEDVTPADRGVIRYVIGVGDAPFSEKSRQELNTIA 304  
DB 246 ARDATKILVITDGEKFGPLGYEDVTPADRGVIRYVIGVGDAPFSEKSRQELNTIA 305

QY 305 SKPPRDVFOVNNPEALKTIQNLREKIFAIEGTQSSSSSFEHMSQEGSAITSNGP 364  
DB 306 SKPSQEHIFKVEDFDALDKDIQTQLREKIFPIEGTETSSSSSFEHMSQEGSAITSNGP 365

QY 365 LLSTVGSYDAGGVFLYTSKEKSTFINMTRVDSMDNAYLGYAAAILLRENVQSLVLCAP 424  
DB 366 VLGAAGFTWEGGAFYPPNNMPTFINNSQENVDMRDSYLGSTELALWKGVQSLVLCAP 425

QY 425 RYQHIGLVAMERQNTKMSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIE 484  
DB 426 RYQHTGRAVITQVSRQRMKAETVGTQISYFGPSLCSVDVDSNGSTDLVLIGAPHYIE 485

QY 485 QTRGGQVSVCPPLRGQARMOCDAVLGYEQQPGWRFGAALTVLGVDNGDKLTDVAIGAP 544  
DB 486 QTRGAQVSVCPPLRGWR-RWMCDAVLGYEQQPGWRFGAALTVLGVDNGDKLTDVAIGAP 544

QY 545 GEENRGAVLPFGTSGSISPSHSORTAGSKLSPLOYFQCSLGGODLTMDGLVDLITV 604  
DB 545 GEENRGAVLPFGTSGSISPSHSORTAGSKLSPLOYFQCSLGGODLTMDGLVDLITV 604

QY 605 GAQGHVLLRSQPLRVLRVKAIMEFNPREVARNVFCNDQVKGKAGEVRLVHVKQSTRD 664  
DB 605 GARGQVLLRLRPVLVWGVSMQFIPABIPRSAPFECREQVSEQLTVQSNICLIDKRSKN 664

QY 665 PLRQGTQSVVYDIALDSCGPHSRVAFNETKSTRCTQVLGTLTQCTETKLQLPNCIE 724  
DB 665 ILGSRDQSSVTLDALDPGLSPRATFOETKNSLSRVRVLGKAHCENFALLPSCVE 724

QY 725 DPVSPVILRLNPSLVGTPLSAFGLNRPVLABDAQLFTALFPFPEKNCNDNICODLSIT 784  
DB 725 DPVSPVILRLNPSLVGTPLSAFGLNRPVLABDAQLFTALFPFPEKNCNDNICODLSIT 784

964 RSLPISLVPLVRLNOTVINDRPOVTFSENLSSTCHTKBELPSHSDFLAELKAPVVC 1023  
DB 963 RDLPSVINFVPELVNQAVWMDVSVHPQNSLCSSEKIAPASDFLAHQKPNVLD 1022

QY 1024 SIACQRIQCDIPFGIOEFNATLKGSLSDWYIKTSHNHLIYSTAEILFNDSVFTLL 1083  
DB 1023 SIAGCLRFCDVPSFSVQBELDFTLKGSLSGFWVRQILQKXSVVSVAEIIFDTSVSQ 1082

QY 1084 PQGAFVSSOTETKVEPEVNPPLSLIVSSVGGILLALITAAALYKLGFFKQVKNMS 1143  
DB 1083 PQGAFVSSOTETKVEPEVNPPLSLIVSSVGGILLALITAAALYKLGFFKQVKNMS 1142

QY 1144 E 1144  
DB 1143 E 1143

RESULT 12

ABU07406  
ID ABU07406 standard; protein; 1163 AA.

AC ABU07406;

DT 28-JAN-2003 (first entry)

DE Protein differentially regulated in prostate cancer #9.

KW Prostate cancer; gene expression; differential regulation;  
KW molecular marker; drug target; cancer detection; cancer diagnosis;  
KW cancer staging; cancer grading; cancer assessing; cancer monitoring.

OS Homo sapiens.

PN W0200281638-A2.

PD 17-OCT-2002.

PF 08-APR-2002; 2002WO-US010824.

PR 06-APR-2001; 2001US-0281731P.

PR 06-APR-2001; 2001US-0281732P.

PA (ORIG-) ORIGENE TECHNOLOGIES INC.

PI Sun Z, Jay G;

XX WP1; 2003-058520/05.

XX Novel genes which are differentially regulated in prostate cancer, useful for diagnosing prostate cancer in prostate tissue sample and assessing therapeutic or preventive intervention in prostate cancer patients.

XX Claim 1; Page 225-228; 416pp; English.

XX The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for the test agent to modulate a biological activity of the polypeptide, and determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating,





Db 604 GARGOVLLLRSLPVLKGVAMRFPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSID 663  
 QY 665 RLREGQIQSVVYDIALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQPNCTE 724  
 Db 664 QL--GDIQSVYRDLALDPLGSLTSAIPNETKPTLTKRKTGLGLHCETLKLPLDCVE 721  
 QY 725 DVPSPVLRNLSVLTPLSAFNLRPVLAEDQRLFTALFPPEKVCNDNICDDLSIT 784  
 Db 722 DVPSPVLRNLSVLTPLSAFNLRPVLAEDQRLFTALFPPEKVCNDNICDDLSIT 781  
 QY 785 FSEMSLDCLVVGGRFNFVTVYRNQDGEDSYRTQVTFPPPLDLSYKRVSTLQNRORSW 844  
 Db 782 LSESGQLTLVGSLSLNLVTVYRNQDGEDSYRTQVTFPPPLDLSYKRVSTLQNRORSW 841  
 QY 845 RLACESASSTEVSGALKSTSCSINHPIFPNSSEVTNITFDVDSKASLGNKULLKANVTS 904  
 Db 842 RLACETV-PTEDRG-LRSSRCSVNHPIFHEGSGNGTFIVTFDVSYKATLGRMLRASASS 899  
 QY 905 ENNMPTNKTEFOLELPVKYAVVMVTSHGVS TKYLPF-TASENTSRVMOCHOVNSNLQ 963  
 Db 900 ENNKASSKATFOLELPVKYAVVMVTSHGVS TKYLPF-TASENTSRVMOCHOVNSNLQ 959  
 QY 964 RSLPISLFLVPLVRLAQTVIWRDPQVTFSENLSSTCHTKERLPSHSDPLAELRKAPVNC 1023  
 Db 960 RDLAISINFVWVLLNGVAVDVMVMEAPSQSL--PCVSEKRPQHSDFLTQISRSPMLDC 1017  
 QY 1024 SIACVORIQCDDIPFGIOBEFNATLKNLSFDWYIKTSHNHLIVTAEILFENDSYVTL 1083  
 Db 1018 SIACVORIQCDDIPFGIOBEFNATLKNLSFDWYIKTSHNHLIVTAEILFENDSYVTL 1077  
 QY 1084 PGQAFVRSOTETKVPFVPEVNPPLTVGSSVGLLILALITAAKYLAFKQYKDMMS 1143  
 Db 1078 PGQAFVRSOTETKVPFVPEVNPPLTVGSSVGLLILALITAAKYLAFKQYKDMMS 1137  
 QY 1144 E 1144  
 Db 1138 D 1138

RESULT 15  
 AAW57491  
 ID AAW57491 standard; protein; 1161 AA.  
 AC AAW57491;  
 XX  
 XX  
 DT 24-AUG-1998 (first entry)  
 XX  
 DE Human beta2 integrin alpha subunit (alpha d) polypeptide.  
 XX  
 KW Beta 2 integrin alpha subunit; alpha d; human; treatment; diabetes;  
 KW reporter-transactivator construct; arteriosclerosis; atherosclerosis;  
 KW inflammatory bowel disease; arthritis; multiple sclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 FH Location/Qualifiers  
 FT 1..16  
 FT /note= "signal peptide"  
 FT 17..1161  
 FT /note= "mature protein"  
 FT 17..1108  
 FT /note= "extracellular domain"  
 FT 150..352  
 FT /note= "I (insertion) domain"  
 FT 1109..1128  
 FT /note= "transmembrane region"  
 FT 1129..1161  
 FT /note= "cytoplasmic domain"  
 XX  
 XX US5766850-A.  
 XX  
 PD 16-JUN-1998.  
 XX

PF 21-DEC-1994; 94US-00362652.  
 XX  
 PR 23-DEC-1993; 93US-00173497.  
 PR 05-AUG-1994; 94US-00286889.  
 XX  
 PA (ICOS-) ICOS CORP.  
 XX  
 XX Van Der Vieren M, Gallatin NM;  
 DR N-PSDB; AAV31540.  
 DR WPI; 1998-361678/31.  
 PT Isolation of DNA encoding protein that binds to integrin subunit - using  
 PT recombinant cells containing reporter-transactivator construct.  
 XX  
 PS Example 5; Col 45-54; 86pp; English.  
 XX  
 CC This represents a human beta2 integrin alpha subunit (alpha d)  
 CC polypeptide. This is used in the methods of the invention for isolating a  
 CC polynucleotide encoding a protein that binds to alpha d. The method  
 CC comprises transforming or transfecting host cells with a DNA construct  
 CC comprising a reporter gene under the control of a promoter regulating a  
 CC transcription factor having a DNA-binding domain and an activating  
 CC domain. A first hybrid DNA sequence encoding a fusion of at least part of  
 CC alpha d and either the DNA-binding domain or the activating domain of the  
 CC transcription factor and a library of hybrid DNA sequences encoding  
 CC fusions of at least part of putative alpha d-binding proteins and the DNA  
 CC binding domain or the activating domain of the transcription factor  
 CC which is not incorporated in the first fusion are expressed in the host  
 CC cells. The binding of an alpha d-binding protein to alpha d in a  
 CC particular host cell is detected by determining production of the  
 CC reporter gene product in the cell. The hybrid DNA sequence encoding the  
 CC alpha d-binding protein can be isolated from the cell. Alpha d may be  
 CC useful for treating graft arteriosclerosis, atherosclerosis, diabetes,  
 CC inflammatory bowel disease, arthritis and multiple sclerosis  
 XX  
 SQ Sequence 1161 AA;  
 Query Match 58.0%; Score 3455; DB 2; Length 1161;  
 Best Local Similarity 59.9%; Pred. No. 7.4e-279;  
 Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;  
 QY 5 VLLLTALTLCHGFNLDTENAMTFQENARGFQSGVYVLOGSRVVGAPQEIIVAAVNRGSLY 64  
 DB 5 VLLLSVLA SYRGNLDVEEPTIFQEDAGFGGSGVYVLOGSRVVGAPQEIIVAAVNRGSLY 64  
 QY 65 QCDYSTGSCETIRLQVPVEAVNMSLGLSLAATTPSPQLLACGPTVHQTCSENTYVKGLCF 124  
 DB 65 DCAATGWCQPIPLHIREAVNMSLGLTLAATNGSRLLACGPTLHRVGENSYSGKSGCL 124  
 QY 125 LFGSNLRQOPQKFPPEALRGCPQEDSDIAFLDGGSGIIPHDFRHKFVSTVMEQLKXSK 184  
 DB 125 LFGSNLRQOPQKFPPEALRGCPQEDSDIAFLDGGSGIIPHDFRHKFVSTVMEQLKXSK 183  
 QY 185 TLPSLMQYSEEPRIHPTFKFONNPNRSLVKPITQLLGRTHATGIRKVVRELFNTING 244  
 DB 184 TLPSLMQYSEEPRIHPTFKFONNPNRSLVKPITQLLGRTHATGIRKVVRELFNTING 243  
 QY 245 ARKNAFKLVITDGEKFGDPLGYEDVTPEDRREGVIRYVIGVGDPAFRSEKRSQELINTIA 304  
 DB 244 ARKNAFKLVITDGEKFGDPLGYEDVTPEDRREGVIRYVIGVGDPAFRSEKRSQELINTIA 303  
 QY 305 SKPPRDHVPQNNFALKTIONQLEKIFATGTGTGSSSSPEHEMSOGFSAATSNCP 364  
 DB 304 SKPPRDHVPQNNFALKTIONQLEKIFATGTGTGSSSSPEHEMSOGFSAATSNCP 363  
 QY 365 LLSTVGSYDVGAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAAIILNRVQSVLVLGAP 424  
 DB 364 LLSTVGSYDVGAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGVAAAIILNRVQSVLVLGAP 423  
 QY 425 RYOHIGLVAMFRONTQWESNANVKGTOIGAYEGSGLCSVDVDSNGSSTLVLIIGAPHYTE 484  
 DB 424 RYOHIGLVAMFRONTQWESNANVKGTOIGAYEGSGLCSVDVDSNGSSTLVLIIGAPHYTE 483



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:16:51 ; Search time 39,4378 Seconds

(without alignments)  
8225.189 Million cell updates/sec

Title: US-09-902-481B-1

Perfect score: 5956

Sequence: 1 MALRVLLLTALTLCHGFNLD.....FKRQYKDMWSEGGPRGAPQ 1153

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 28138677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*

```
1: /cgn2_6/ptodata1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata1/pubpaa/PCT_NEW PUB.pep.*
3: /cgn2_6/ptodata1/pubpaa/US06_NEW PUB.pep.*
4: /cgn2_6/ptodata1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata1/pubpaa/US07_NEW PUB.pep.*
6: /cgn2_6/ptodata1/pubpaa/PCTU5_PUBCOMB.pep.*
7: /cgn2_6/ptodata1/pubpaa/US08_NEW PUB.pep.*
8: /cgn2_6/ptodata1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata1/pubpaa/US09_NEW PUB.pep.*
13: /cgn2_6/ptodata1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata1/pubpaa/US10_NEW PUB.pep.*
17: /cgn2_6/ptodata1/pubpaa/US60_NEW PUB.pep.*
18: /cgn2_6/ptodata1/pubpaa/US60_PUBCOMB.pep.*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5956	100.0	1153	9 US-09-350-259-3	Sequence 3, Appli
2	5956	100.0	1153	10 US-09-902-481A-1	Sequence 1, Appli
3	5956	100.0	1153	10 US-09-891-943-3	Sequence 3, Appli
4	5956	100.0	1153	14 US-10-144-259-30	Sequence 30, Appl
5	5956	100.0	1153	14 US-10-207-655-176	Sequence 176, App
6	5940.5	99.7	1152	9 US-09-945-265-4	Sequence 4, Appli
7	5868	98.5	1137	10 US-09-902-481A-6	Sequence 6, Appli
8	5862	98.4	1137	10 US-09-902-481A-5	Sequence 5, Appli
9	5852	98.3	1137	10 US-09-902-481A-4	Sequence 4, Appli
10	5839	98.0	1137	10 US-09-902-481A-3	Sequence 3, Appli
11	3503	58.8	1163	15 US-10-116-275-204	Sequence 204, App
12	3480	58.4	1163	9 US-09-350-259-4	Sequence 4, Appli
13	3480	58.4	1163	10 US-09-891-943-4	Sequence 4, Appli
14	3455	58.0	1161	10 US-09-350-259-2	Sequence 2, Appli
15	3455	58.0	1161	10 US-09-891-943-2	Sequence 2, Appli

```
16 3439.5 57.7 1161 9 US-09-350-259-99
17 3439.5 57.7 1161 10 US-09-891-943-99
18 3270.5 54.9 1161 9 US-09-350-259-55
19 3270.5 54.9 1161 10 US-09-891-943-55
20 3264 54.8 1161 9 US-09-350-259-53
21 3264 54.8 1161 10 US-09-891-943-53
22 3251.5 54.6 1151 9 US-09-350-259-37
23 3251.5 54.6 1151 10 US-09-891-943-37
24 3243 54.4 1155 9 US-09-350-259-46
25 3243 54.4 1155 10 US-09-891-943-46
26 1848 31.0 369 12 US-10-087-192-1212
27 1563 26.2 1170 9 US-09-945-265-2
28 1559 26.2 1170 12 US-10-261-164-1
29 1535.5 25.8 1223 16 US-10-408-765A-295
30 1361 22.9 1086 16 US-10-408-765A-1871
31 1229.5 20.6 494 9 US-09-350-259-103
32 1229.5 20.6 494 10 US-09-891-943-103
33 1189.5 20.0 413 9 US-09-350-259-101
34 1189.5 20.0 413 10 US-09-891-943-101
35 1163 19.5 1179 14 US-10-177-550-2
36 1163 19.5 1179 15 US-10-173-551-2
37 1123.5 18.9 1179 12 US-09-918-715-250
38 1110 18.6 1188 15 US-10-291-265-810
39 1108.5 18.6 1151 10 US-09-984-130-103
40 1108.5 18.6 1151 10 US-09-836-353A-103
41 1106 18.6 1188 15 US-10-291-265-338
42 1102.5 18.5 1189 10 US-09-984-130-35
43 1102.5 18.5 1189 10 US-09-836-353A-35
44 1102.5 18.5 1189 12 US-10-262-839-4
45 1087.5 18.3 1181 12 US-10-211-462-187
```

#### ALIGNMENTS

#### RESULT 1

```
US-09-350-259-3
; Sequence 3, Application US/09350259
; Parent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3
```

Query Match 100.0% Score 5956; DB 9; Length 1153;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MALRVLLLTALTLCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEIIVAAQR 60

DB 1 MALRVLLLTALTLCHGFNLDTENAMTFQENARGFQSVVQLQSGRVVVGAPQEIIVAAQR 60

OY 61 GSLVQCDYSTGSCPIRLQVPEAVNMISGLSLAATTPPQLLACGPTVHTQCSNTYVK 120

Db	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPFQLLACQFTVQTCSENYYK	120
Qy	121	GLCFLFGSNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL	180
Db	121	GLCFLFGSNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL	180
Qy	181	KKSKTLFSLMOYSEEFRIHPTFKFQNNPNRSLVKPTIQLLGRTHATGIRKVVRELFN	240
Db	181	KKSKTLFSLMOYSEEFRIHPTFKFQNNPNRSLVKPTIQLLGRTHATGIRKVVRELFN	240
Qy	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQEL	300
Db	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQEL	300
Qy	301	NTIASKPPDRHVQVNNPEALKTIQNLREKI PAIEGTQTGSSSSFEHMSQEGFSAAIT	360
Db	301	NTIASKPPDRHVQVNNPEALKTIQNLREKI PAIEGTQTGSSSSFEHMSQEGFSAAIT	360
Qy	361	SNGPLLSTVGSYDWAQGVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV	420
Db	361	SNGPLLSTVGSYDWAQGVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV	420
Qy	421	LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP	480
Db	421	LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP	480
Qy	481	HYTEQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDA	540
Db	481	HYTEQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDA	540
Qy	541	IGAPGEEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV	600
Db	541	IGAPGEEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV	600

Db	1141	NMSEGGPPCAEPQ	1153
RESULT 2			
US-09-902-481A-1			
; Sequence 1, Application US/09902481A			
; Publication No. US20030054440A1			
; GENERAL INFORMATION:			
; APPLICANT: Springer, Timothy			
; APPLICANT: Shimaoka, Motomu			
; APPLICANT: Shifman, Julia			
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY			
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK			
; CURRENT APPLICATION NUMBER: US/09/902,481A			
; CURRENT FILING DATE: 2001-07-09			
; PRIOR APPLICATION NUMBER: US 60/216,600			
; PRIOR FILING DATE: 2000-07-07			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: Patent in version 3.1			
; SEQ ID NO 1			
; LENGTH: 1153			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: mat_peptide			
; LOCATION: (17)..()			
; OTHER INFORMATION:			
US-09-902-481A-1			
Query Match 100.0%; Score 5956; DB 10; Length 1153;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MALRVLLTALTLCCHGFNLDTENAMTFOENARGFQSVVQLQGSRVVVVGAPOEIVAAQR	60
Db	1	MALRVLLTALTLCCHGFNLDTENAMTFOENARGFQSVVQLQGSRVVVVGAPOEIVAAQR	60
Qy	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPFQLLACQFTVQTCSENYYK	120
Db	61	GSLYQCDYSTGCEPIRLQVPEAVNMSLGLSLAATTSPFQLLACQFTVQTCSENYYK	120
Qy	121	GLCFLFGSNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL	180
Db	121	GLCFLFGSNLRQOPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQL	180
Qy	181	KKSKTLFSLMOYSEEFRIHPTFKFQNNPNRSLVKPTIQLLGRTHATGIRKVVRELFN	240
Db	181	KKSKTLFSLMOYSEEFRIHPTFKFQNNPNRSLVKPTIQLLGRTHATGIRKVVRELFN	240
Qy	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQEL	300
Db	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQEL	300
Qy	301	NTIASKPPDRHVQVNNPEALKTIQNLREKI PAIEGTQTGSSSSFEHMSQEGFSAAIT	360
Db	301	NTIASKPPDRHVQVNNPEALKTIQNLREKI PAIEGTQTGSSSSFEHMSQEGFSAAIT	360
Qy	361	SNGPLLSTVGSYDWAQGVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV	420
Db	361	SNGPLLSTVGSYDWAQGVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV	420
Qy	421	LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP	480
Db	421	LGAPRYQHIGLVAMFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP	480
Qy	481	HYTEQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDA	540
Db	481	HYTEQTRGGQVSVCPPLRGQRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDA	540
Qy	541	IGAPGEEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV	600
Db	541	IGAPGEEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV	600

QY 601 DLTGAGHVLRLRSQVLRVKAIMEFNPREFNARVNFECNDQVVKGEAGVRVCLHVOK 660  
DB 601 DLTGAGHVLRLRSQVLRVKAIMEFNPREFNARVNFECNDQVVKGEAGVRVCLHVOK 660  
QY 661 STRDLREGQIQSVVYDLDLSDGRPHSRVAVNFNETNSTRTQVGLTQTCETLKLQLP 720  
DB 661 STRDLREGQIQSVVYDLDLSDGRPHSRVAVNFNETNSTRTQVGLTQTCETLKLQLP 720  
QY 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRPVLAEDAQRFTALFPPEKNCNDNICODD 780  
DB 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRPVLAEDAQRFTALFPPEKNCNDNICODD 780  
QY 781 LSITPFSMSLDCLVVGGRPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840  
DB 781 LSITPFSMSLDCLVVGGRPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840  
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASIGNKLLKA 900  
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASIGNKLLKA 900  
QY 901 NVTSENMPRTNKTETFOLELPVKYAVVMVTSHGVSSTKYLNFNTASENTSRVMQHQYQVSN 960  
DB 901 NVTSENMPRTNKTETFOLELPVKYAVVMVTSHGVSSTKYLNFNTASENTSRVMQHQYQVSN 960  
QY 961 LGQSLPISLVFLVPLVRLNQTVIWDREPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020  
DB 961 LGQSLPISLVFLVPLVRLNQTVIWDREPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020  
QY 1021 VNCSTAVCQRIQCDIPFGIQEENFATLKGMLSPDWYIKTSHNLLIIVSTAEIIFNDSVF 1080  
DB 1021 VNCSTAVCQRIQCDIPFGIQEENFATLKGMLSPDWYIKTSHNLLIIVSTAEIIFNDSVF 1080  
QY 1081 TLLPQGGAFVRSQETKVEPPEFVNPPLPLIVGSSVGGILLALITAAALYKLGFFKRYKD 1140  
DB 1081 TLLPQGGAFVRSQETKVEPPEFVNPPLPLIVGSSVGGILLALITAAALYKLGFFKRYKD 1140  
QY 1141 MMSBGGPPGABEQ 1153  
DB 1141 MMSBGGPPGABEQ 1153

RESULT 3  
US-09-891-943-3  
; Sequence 3, Application US/09891943  
; Publication No. US2003007728A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US2003007728A1 Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/891,943  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-891-943-3

Query Match 100.0%; Score 5956; DB 10; Length 1153;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTILCHGFNLDTENAMTFQENARFGQSVVQLQGSRRVVVGPAQBIIVANQR 60  
DB 1 MALRVLLLTALTILCHGFNLDTENAMTFQENARFGQSVVQLQGSRRVVVGPAQBIIVANQR 60  
QY 61 GSIYQCDYSTSCCEPRLQVPVEAVNMSLCESLAATTSPPOLLACGPTVHTQTCSENIVYK 120  
DB 61 GSIYQCDYSTSCCEPRLQVPVEAVNMSLCESLAATTSPPOLLACGPTVHTQTCSENIVYK 120  
QY 121 GLCFLFGSNLRQCPQKPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQL 180  
DB 121 GLCFLFGSNLRQCPQKPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQL 180  
QY 181 KSKTILFSLMOYSEEFRIHFTFKFQNNPNRPSIAVKPITOLLGRTHTATGIRKVVRLFN 240  
DB 181 KSKTILFSLMOYSEEFRIHFTFKFQNNPNRPSIAVKPITOLLGRTHTATGIRKVVRLFN 240  
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGYGDFAFRSEKSOEL 300  
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVIPEADREGVIRYVIGYGDFAFRSEKSOEL 300  
QY 301 NTIASKPRPHVQVNNFEALKTIQNLREKIPAIETQTCSSSSFEHMSQEGFSAIT 360  
DB 301 NTIASKPRPHVQVNNFEALKTIQNLREKIPAIETQTCSSSSFEHMSQEGFSAIT 360  
QY 361 SNGELLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIIRNRVQSLV 420  
DB 361 SNGELLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIIRNRVQSLV 420  
QY 421 LGAPRYQHI GLVAMFRONTGWESNANVKQTQIGAFGASLCSVDVDSNGSTDLVLGAP 480  
DB 421 LGAPRYQHI GLVAMFRONTGWESNANVKQTQIGAFGASLCSVDVDSNGSTDLVLGAP 480  
QY 481 HYBTQTRGGQVSVCLPRGQARWQCDVLYGEGQPGWGFMAALTVLGDVNGDKLTDVA 540  
DB 481 HYBTQTRGGQVSVCLPRGQARWQCDVLYGEGQPGWGFMAALTVLGDVNGDKLTDVA 540  
QY 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSRPLQVPGQSLSGQDITMDGLV 600  
DB 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSRPLQVPGQSLSGQDITMDGLV 600  
QY 601 DLTGAGHVLRLRSQVLRVKAIMEFNPREFNARVNFECNDQVVKGEAGVRVCLHVOK 660  
DB 601 DLTGAGHVLRLRSQVLRVKAIMEFNPREFNARVNFECNDQVVKGEAGVRVCLHVOK 660  
QY 661 STRDLREGQIQSVVYDLDLSDGRPHSRVAVNFNETNSTRTQVGLTQTCETLKLQLP 720  
DB 661 STRDLREGQIQSVVYDLDLSDGRPHSRVAVNFNETNSTRTQVGLTQTCETLKLQLP 720  
QY 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRPVLAEDAQRFTALFPPEKNCNDNICODD 780  
DB 721 NCIEDPVSPIVLRNFSLVGTPLSAFGNLRPVLAEDAQRFTALFPPEKNCNDNICODD 780  
QY 781 LSITPFSMSLDCLVVGGRPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840  
DB 781 LSITPFSMSLDCLVVGGRPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840  
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASIGNKLLKA 900  
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASIGNKLLKA 900  
QY 901 NVTSENMPRTNKTETFOLELPVKYAVVMVTSHGVSSTKYLNFNTASENTSRVMQHQYQVSN 960  
DB 901 NVTSENMPRTNKTETFOLELPVKYAVVMVTSHGVSSTKYLNFNTASENTSRVMQHQYQVSN 960  
QY 961 LGQSLPISLVFLVPLVRLNQTVIWDREPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020  
DB 961 LGQSLPISLVFLVPLVRLNQTVIWDREPQVTFSENLSTCHTKERLPSHSDFLAELRKAPV 1020  
QY 1021 VNCSTAVCQRIQCDIPFGIQEENFATLKGMLSPDWYIKTSHNLLIIVSTAEIIFNDSVF 1080  
DB 1021 VNCSTAVCQRIQCDIPFGIQEENFATLKGMLSPDWYIKTSHNLLIIVSTAEIIFNDSVF 1080  
QY 1081 TLLPQGGAFVRSQETKVEPPEFVNPPLPLIVGSSVGGILLALITAAALYKLGFFKRYKD 1140

```
||||| 1081 TLLPGQAGVRSQTEKVEPFEVNEPLTVGSSVGGLLALLLTAALYKLGFFKQYKD 1140
||||| 1141 MMSEGGPPGAEPPQ 1153
||||| 1141 MMSEGGPPGAEPPQ 1153

RESULT 4
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30

Query Match 100.0%; Score 5956; DB 14; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLOGSRVVVGAPOBIVAANOR 60
DB 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLOGSRVVVGAPOBIVAANOR 60
QY 61 GSYVQCDYSGSCPIRLQVPEAVNMSIGLSLAATTSPPLLACGPTVHQTCSNTYVK 120
DB 61 GSYVQCDYSGSCPIRLQVPEAVNMSIGLSLAATTSPPLLACGPTVHQTCSNTYVK 120
QY 121 GLCFLGSLNRQPKQPFPEALRGCPQEDSDIAPLIDGSGSIIPHPDFRMKKEFVSTVNEQL 180
DB 121 GLCFLGSLNRQPKQPFPEALRGCPQEDSDIAPLIDGSGSIIPHPDFRMKKEFVSTVNEQL 180
QY 181 KKSMTLPSLMQYSEEPRIHETKEFQNNPNRSLVXPDITOLLGRTHTATGIRKVVRELFN 240
DB 181 KKSMTLPSLMQYSEEPRIHETKEFQNNPNRSLVXPDITOLLGRTHTATGIRKVVRELFN 240
QY 241 ITNGARKNAFKILVVIITDGKFGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSRQBL 300
DB 241 ITNGARKNAFKILVVIITDGKFGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSRQBL 300
QY 301 NTIASKPPRHHVQVNNFALKTIQNLREKFAIEGTQGTSSSEHEMSQGFSAAIT 360
DB 301 NTIASKPPRHHVQVNNFALKTIQNLREKFAIEGTQGTSSSEHEMSQGFSAAIT 360
QY 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLV 420
DB 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLV 420
QY 421 LGAPRYQHIGLVAMFRONTGWESNANVKTQIGAYFGASLCSDVDVDSNGSTDLVLIGAP 480
DB 421 LGAPRYQHIGLVAMFRONTGWESNANVKTQIGAYFGASLCSDVDVDSNGSTDLVLIGAP 480
QY 481 HYTBOTGGGVSVCPPLPRGQARQWCDVILYGEQGWGFGAALTIVLGDVNGDKLTDVA 540
DB 481 HYTBOTGGGVSVCPPLPRGQARQWCDVILYGEQGWGFGAALTIVLGDVNGDKLTDVA 540
QY 541 IGAPGEDBNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
```

```
||||| 541 IGAPGEDBNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600
||||| 601 DLTVGAQGHVLLRSQPVLRVKAIIMEPNPREVARNVFECDNVVKGKEAGEVVRVCLHVQK 660
||||| 601 DLTVGAQGHVLLRSQPVLRVKAIIMEPNPREVARNVFECDNVVKGKEAGEVVRVCLHVQK 660
||||| 661 STROBLREGQIOSVVTYDLALDSGRPHSRVAFNETKNSSTRETQVGLGTQTCETLKLQLP 720
||||| 661 STROBLREGQIOSVVTYDLALDSGRPHSRVAFNETKNSSTRETQVGLGTQTCETLKLQLP 720
||||| 721 NCIEDPVSPVILRLNPLSLVGTPLSAFGLNLRPLVLAEDAQRLEFALFPPEKNCNDNICQDD 780
||||| 721 NCIEDPVSPVILRLNPLSLVGTPLSAFGLNLRPLVLAEDAQRLEFALFPPEKNCNDNICQDD 780
||||| 781 LSITPSFMSLCLVVGGBREBNVTVVANDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840
||||| 781 LSITPSFMSLCLVVGGBREBNVTVVANDGEDSYRTQVTFPPFLDLSYRKVSTLQORS 840
||||| 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
||||| 841 QRSWLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKA 900
||||| 901 NVTSENMPRNKTEFQLELPVKYAVYVWVTSHGVSSTKYLNFASNTSRVMOHQYQVSN 960
||||| 901 NVTSENMPRNKTEFQLELPVKYAVYVWVTSHGVSSTKYLNFASNTSRVMOHQYQVSN 960
||||| 961 LGORSPLPSLVPLVPLNQTIVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
||||| 961 LGORSPLPSLVPLVPLNQTIVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020
||||| 1021 VNCSTIACVQRIQCDIPFFGIOEBFNATLKGNSLSDWYIKTSHNHLIIIVSTAILFNDVSVF 1080
||||| 1021 VNCSTIACVQRIQCDIPFFGIOEBFNATLKGNSLSDWYIKTSHNHLIIIVSTAILFNDVSVF 1080
||||| 1081 TLLPGQAGVRSQTEKVEPFEVNEPLTVGSSVGGLLALLLTAALYKLGFFKQYKD 1140
||||| 1081 TLLPGQAGVRSQTEKVEPFEVNEPLTVGSSVGGLLALLLTAALYKLGFFKQYKD 1140
||||| 1141 MMSEGGPPGAEPPQ 1153
||||| 1141 MMSEGGPPGAEPPQ 1153

RESULT 5
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 100.0%; Score 5956; DB 14; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLOGSRVVVGAPOBIVAANOR 60
DB 1 MALRVLLLTALTLCGHNLDENAMTFQENARGFGQSVVQLOGSRVVVGAPOBIVAANOR 60
QY 61 GSYVQCDYSGSCPIRLQVPEAVNMSIGLSLAATTSPPLLACGPTVHQTCSNTYVK 120
```

```
Db 61 GSIYQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120
Qy 121 GLCFPLGSLNRQOPKPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Db 121 GLCFPLGSLNRQOPKPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180
Qy 181 KSKKTLFSLMQYSEEFRIHFTFKFQONNPNPSLVKPIITQLGRTHATGIRKVVRELFN 240
Db 181 KSKKTLFSLMQYSEEFRIHFTFKFQONNPNPSLVKPIITQLGRTHATGIRKVVRELFN 240
Qy 241 ITNGARKNAFKILVITDGEKGDPLGYEDVTPADREGVIRYVIGVDGDAFRSEKSRQEL 300
Db 241 ITNGARKNAFKILVITDGEKGDPLGYEDVTPADREGVIRYVIGVDGDAFRSEKSRQEL 300
Qy 301 NTIASKPPDRHVFQVNNFEALKTIONQIREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360
Db 301 NTIASKPPDRHVFQVNNFEALKTIONQIREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360
Qy 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLV 420
Db 361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLV 420
Qy 421 LGAPRYQHIGLVAMPFRONTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Db 421 LGAPRYQHIGLVAMPFRONTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDVLIGAP 480
Qy 481 HYETQTRGQVSVCPILPRGORARWQCDVLYGEGQOPWGRFGAALFVLGDNVNGDKLTVA 540
Db 481 HYETQTRGQVSVCPILPRGORARWQCDVLYGEGQOPWGRFGAALFVLGDNVNGDKLTVA 540
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600
Qy 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVOK 660
Db 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVOK 660
Qy 720 STRDRUREGQISQVVTYDALDSGRPHSRAVNTKSTNRTOVLGLTQTCETLKLQLP 720
Db 720 STRDRUREGQISQVVTYDALDSGRPHSRAVNTKSTNRTOVLGLTQTCETLKLQLP 720
Qy 721 NCIEDPVPISVLRLNLSLWGTPLSAFENLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
Db 721 NCIEDPVPISVLRLNLSLWGTPLSAFENLRPVLAEDAQRLLFTALPFPEKNCNDNICQDD 780
Qy 781 LSITPFSMLDCLVWGGPREFNVTVTRNDGDSYRTQVTFPPLDLSYRKVSTLQNRS 840
Db 781 LSITPFSMLDCLVWGGPREFNVTVTRNDGDSYRTQVTFPPLDLSYRKVSTLQNRS 840
Qy 841 QRSWRLACBSASSTEVSGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKA 900
Db 841 QRSWRLACBSASSTEVSGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKA 900
Qy 901 NVTSNNMPRTNKTETFEQLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMOHQVQSN 960
Db 901 NVTSNNMPRTNKTETFEQLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMOHQVQSN 960
Qy 961 LQORSPLISLVFLVPLVRLNQTWMDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020
Db 961 LQORSPLISLVFLVPLVRLNQTWMDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPV 1020
Qy 1021 VNCISIAVCQRIQCDIIPFGIQEDEFNATLKGNLSDFWYIKTSHNLLIVSTABILFNDSPV 1080
Db 1021 VNCISIAVCQRIQCDIIPFGIQEDEFNATLKGNLSDFWYIKTSHNLLIVSTABILFNDSPV 1080
Qy 1081 TLLPGQAFVRSQTEKVPFEPVNPFLPLIVGSSVGGHLLALLITAAALKYKGFKKOYKD 1140
Db 1081 TLLPGQAFVRSQTEKVPFEPVNPFLPLIVGSSVGGHLLALLITAAALKYKGFKKOYKD 1140
Qy 1141 MWSEGGPPGAEPQ 1153
Db 1141 MWSEGGPPGAEPQ 1153
```

## RESULT 5

```
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: CN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4
```

Query Match 99.7%; Score 5940.5; DB 9; Length 1152;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1152; Conservative 0; Mismatches

0; Indels 1; Gaps 1;

1 MALRVLLLTALTTLCHGFNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQEIIVAAANOR 60

1 MALRVLLLTALTTLCHGFNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQEIIVAAANOR 60

61 GSLVQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120

61 GSLVQCDYSTGSCPIRLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVK 120

121 GLCFPLGSLNRQOPKPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180

121 GLCFPLGSLNRQOPKPEALRGCGQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180

181 KSKKTLFSLMQYSEEFRIHFTFKFQONNPNPSLVKPIITQLGRTHATGIRKVVRELFN 240

181 KSKKTLFSLMQYSEEFRIHFTFKFQONNPNPSLVKPIITQLGRTHATGIRKVVRELFN 240

241 ITNGARKNAFKILVITDGEKGDPLGYEDVTPADREGVIRYVIGVDGDAFRSEKSRQEL 300

241 ITNGARKNAFKILVITDGEKGDPLGYEDVTPADREGVIRYVIGVDGDAFRSEKSRQEL 300

301 NTIASKPPDRHVFQVNNFEALKTIONQIREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360

301 NTIASKPPDRHVFQVNNFEALKTIONQIREKIFAIEGTQTGSSSFHEHMSQEGFSAIT 360

361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLV 420

361 SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLV 420

421 LGAPRYQHIGLVAMPFRONTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDVLIGAP 480

421 LGAPRYQHIGLVAMPFRONTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDVLIGAP 480

481 HYETQTRGQVSVCPILPRGORARWQCDVLYGEGQOPWGRFGAALFVLGDNVNGDKLTVA 540

481 HYETQTRGQVSVCPILPRGORARWQCDVLYGEGQOPWGRFGAALFVLGDNVNGDKLTVA 540

541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600

541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLV 600

601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVOK 660

601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECDQVWVKEAGEVRVCLHVOK 660

QY 661 STDRLEBQIQSVVTVYDLALDSGRHSRAVFNETKNSRRQTVGLTQTCETLKLQLP 720  
DB 660 STDRLEBQIQSVVTVYDLALDSGRHSRAVFNETKNSRRQTVGLTQTCETLKLQLP 719  
QY 721 NCIEDPVSPIVLRNLSVGLTSLAFNGLRPVLAEDAQRLEFTALPFPFKXCGNDNICQDD 780  
DB 720 NCIEDPVSPIVLRNLSVGLTSLAFNGLRPVLAEDAQRLEFTALPFPFKXCGNDNICQDD 779  
QY 781 LSITFSMSLDCLVVGCPREFNVTVVRNDGEDSYRTQVTFPEPLDLSYRKVSTLQONS 840  
DB 780 LSITFSMSLDCLVVGCPREFNVTVVRNDGEDSYRTQVTFPEPLDLSYRKVSTLQONS 839  
QY 841 QRSWRLACESASSTEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKA 900  
DB 840 QRSWRLACESASSTEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKA 899  
QY 901 NVTSENMPRTNKTETQLELPVKYAVYVMTVSHGVSTKYLNFPTASNTSRVMOHOYQVSN 960  
DB 900 NVTSENMPRTNKTETQLELPVKYAVYVMTVSHGVSTKYLNFPTASNTSRVMOHOYQVSN 959  
QY 961 LGORSLSPISLVFLVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA 1020  
DB 960 LGORSLSPISLVFLVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA 1019  
QY 1021 VNCISIAVCQRIQCDIPFPFGIOBENATLKNLSFDWYIKTSHNHLIIVSTABILFNDSP 1080  
DB 1020 VNCISIAVCQRIQCDIPFPFGIOBENATLKNLSFDWYIKTSHNHLIIVSTABILFNDSP 1079  
QY 1081 TLLPGQGAFFRSOTETKVEPPEVFNPLPLIVGSSVGGLLMLLALITAAALYKLGFFKYOYKD 1140  
DB 1080 TLLPGQGAFFRSOTETKVEPPEVFNPLPLIVGSSVGGLLMLLALITAAALYKLGFFKYOYKD 1139  
QY 1141 MMSEGGPPGAEPQ 1153  
DB 1140 MMSEGGPPGAEPQ 1152

RESULT 7  
US-09-902-481a-6  
; Sequence 6, Application US/09902481a  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimooka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902.481a  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 6  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481a-6

Query Match 98.5%; Score 5868; DB 10; Length 1137;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNLTENAMTQENARGQSVVOLQSRVTVVVGAPQBEIVAAORGSYQCDYSTGSCPEI 76  
DB 1 FNLTENAMTQENARGQSVVOLQSRVTVVVGAPQBEIVAAORGSYQCDYSTGSCPEI 60

QY 77 RLOPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQEQFK 136  
DB 61 RLOPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQEQFK 120

QY 137 PPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTILFSLMOYSBEF 196  
DB 136 PPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTILFSLMOYSBEF 180  
QY 197 RIHFTFKFQNNPNRSLVKIPITQLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256  
DB 196 RIHFTFKFQNNPNRSLVKIPITQLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 240  
QY 257 TDGKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDRHVFQV 316  
DB 256 TDGKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDRHVFQV 300  
QY 317 NFEALKTIQONLREKIPAIETGTTGSSSSFEHKSQEGFSAAITNSNGPILLSVGSYDAG 376  
DB 316 NFEALKTIQONLREKIPAIETGTTGSSSSFEHKSQEGFSAAITNSNGPILLSVGSYDAG 360  
QY 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVLAGAPRYOHIGLVAMFR 436  
DB 376 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVLAGAPRYOHIGLVAMFR 420  
QY 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYBQTRGGQSVCP 496  
DB 436 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYBQTRGGQSVCP 480  
QY 497 PRGQBARWQCDVLYGEGQPGWGRGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 556  
DB 496 PRGQBARWQCDVLYGEGQPGWGRGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 540  
QY 557 HGTSGSGISPSHSOSIAGSKLSPRLOYFGQSLSGQDQDITMDGLVLTGVAQGHVILLASQ 616  
DB 556 HGTSGSGISPSHSOSIAGSKLSPRLOYFGQSLSGQDQDITMDGLVLTGVAQGHVILLASQ 600  
QY 617 PVLRVKAIMERNPREVARNVPECNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 676  
DB 616 PVLRVKAIMERNPREVARNVPECNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 660  
QY 677 YDLALDSGRPHSRVFNETKNSRRQTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 736  
DB 676 YDLALDSGRPHSRVFNETKNSRRQTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720  
QY 737 SLVGTPLSAFAGNLRPVLAEADAORLFTALPFPFKXCGNDNICODDLSITPFSNLSCLVVG 796  
DB 736 SLVGTPLSAFAGNLRPVLAEADAORLFTALPFPFKXCGNDNICODDLSITPFSNLSCLVVG 780  
QY 797 GPREFNVTVVRNDGEDSYRTQVTFPEPLDLSYRKVSTLQONSORSMRLACESASSTEV 856  
DB 796 GPREFNVTVVRNDGEDSYRTQVTFPEPLDLSYRKVSTLQONSORSMRLACESASSTEV 840  
QY 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKXANVTSENMPRTNKTBF 916  
DB 856 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKXANVTSENMPRTNKTBF 900  
QY 917 QLELPVKYAVYVMTVSHGVSTKYLNFPTASNTSRVMOHOYQVSNLQORSLSPLVFLPV 976  
DB 916 QLELPVKYAVYVMTVSHGVSTKYLNFPTASNTSRVMOHOYQVSNLQORSLSPLVFLPV 960  
QY 977 RLNQTVIWDPRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVVNCISIAVCQRIQCDIP 1036  
DB 976 RLNQTVIWDPRPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PVVNCISIAVCQRIQCDIP 1020  
QY 1037 FPGIOBENATLKNLSFDWYIKTSHNHLIIVSTABILFNDSPVTLPGQGAFFRSOTET 1096  
DB 1036 FPGIOBENATLKNLSFDWYIKTSHNHLIIVSTABILFNDSPVTLPGQGAFFRSOTET 1080  
QY 1097 KVEPPEVFNPLPLIVGSSVGGLLMLLALITAAALYKLGFFKYOYKDMSEGGPPGAEPQ 1153  
DB 1096 KVEPPEVFNPLPLIVGSSVGGLLMLLALITAAALYKLGFFKYOYKDMSEGGPPGAEPQ 1137

RESULT 8  
US-09-902-481a-5  
; Sequence 5, Application US/09902481a

Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimooka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-5

Query Match 98.4%; Score 5862; DB 10; Length 1137;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76  
Db 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTSSENTYVKGCLFGLFGLNLRQOQPK 136  
Db 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTSSENTYVKGCLFGLFGLNLRQOQPK 120

QY 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTWMEQLKSKTLFSLMOYSEEF 196  
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTWMEQLKSKTLFSLMOYSEEF 180

QY 197 RIHFTKFEQNNPNSRSLVKTQLLGRTHATGIRKVVRELFTNGARKNAFKILVVI 256  
Db 181 RIHFTKFEQNNPNSRSLVKTQLLGRTHATGIRKVVRELFTNGARKNAFKILVVI 240

QY 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVGDFAFRSEKSRQELNTIASKPPRDHVFQV 316  
Db 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVGDFAFRSEKSRQELNTIASKPPRDHVFQV 300

QY 317 NFEALKTIONLREKIFAIEGHTQSSSSFEHMSQEGPSAAITNGPLLSITVGSYDRAG 376  
Db 301 NFEALKTIONLREKIFAIEGHTQSSSSFEHMSQEGPSAAITNGPLLSITVGSYDRAG 360

QY 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNRRVQSLVGLGAPRYQHIGLVAMFR 436  
Db 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNRRVQSLVGLGAPRYQHIGLVAMFR 420

QY 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAPHYETRGQGVSVCP 496  
Db 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAPHYETRGQGVSVCP 480

QY 497 PRGQARWOCDAVLVGEQGPWGRFGAALTVDLVNNGDKLTDVAICAGEEDNRGAVYLF 556  
Db 481 PRGQARWOCDAVLVGEQGPWGRFGAALTVDLVNNGDKLTDVAICAGEEDNRGAVYLF 540

QY 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGGAQGHVILLRSQ 616  
Db 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGGAQGHVILLRSQ 600

QY 617 PVLRVKALINEFNPREVARNVFCNDQVVKGEAGEVRVCLHWOKSTRDLREGQIQSVVT 676  
Db 601 PVLRVKALINEFNPREVARNVFCNDQVVKGEAGEVRVCLHWOKSTRDLREGQIQSVVT 660

QY 677 YDLALDSGRPHSAVNETKNSRTRQTQVLGLTQTCSTLKLQLPNCIEDPVSPIVRLNPF 736  
Db 661 YDLALDSGRPHSAVNETKNSRTRQTQVLGLTQTCSTLKLQLPNCIEDPVSPIVRLNPF 720

QY 737 SLVGTPLSAFGNLRPVLAEDAQRLEFTALFFPEKNCNDNQCDDLSITTFPMSLDCLVVG 796  
Db 721 SLVGTPLSAFGNLRPVLAEDAQRLEFTALFFPEKNCNDNQCDDLSITTFPMSLDCLVVG 780

QY 797 GPREFNVTVTRNDGEDSVETQVTFPEPLDLSYKXVSTLQNSQSRWRLACBSASSTEV 856  
Db 781 GPREFNVTVTRNDGEDSVETQVTFPEPLDLSYKXVSTLQNSQSRWRLACBSASSTEV 840

QY 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKXANTSENMPRTNKTET 916  
Db 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKXANTSENMPRTNKTET 900

QY 917 QLELPVKIAYTMVTVSHGVSTKYLNFITASENTSRVMOHOYOVSNLQORSIPISLVLVLPV 976  
Db 901 QLELPVKIAYTMVTVSHGVSTKYLNFITASENTSRVMOHOYOVSNLQORSIPISLVLVLPV 960

QY 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDFLAELRKA PVVNCIAVCQRIQCDIP 1036  
Db 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDFLAELRKA PVVNCIAVCQRIQCDIP 1020

QY 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQGFVRSQTET 1096  
Db 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQGFVRSQTET 1080

QY 1097 KVEPFEVNPPLPIVGSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPPQ 1153  
Db 1081 KVEPFEVNPPLPIVGSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPPQ 1137

RESULT 9  
US-09-902-481A-4  
; Sequence 4: Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimooka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-4

Query Match 98.3%; Score 5852; DB 10; Length 1137;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76  
Db 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60

QY 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTSSENTYVKGCLFGLFGLNLRQOQPK 136  
Db 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTSSENTYVKGCLFGLFGLNLRQOQPK 120

QY 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTWMEQLKSKTLFSLMOYSEEF 196  
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTWMEQLKSKTLFSLMOYSEEF 180

QY 197 RIHFTKFEQNNPNSRSLVKTQLLGRTHATGIRKVVRELFTNGARKNAFKILVVI 256  
Db 181 RIHFTKFEQNNPNSRSLVKTQLLGRTHATGIRKVVRELFTNGARKNAFKILVVI 240

Db 181 RIHFTFZEFQWNNPRSLIKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240  
Qy 257 TDGKFGDGLYEDVPEADREGVIRVVGIGDAFSEKSRQELNTIASRPHDHFQVN 316  
Db 241 TDGKFGDGLYEDVPEADREGVIRVVGIGDAFSEKSRQELNTIASRPHDHFQVN 300  
Qy 317 NFELKTIQNLREKIFAIETGQTGSSSSSEHEMSQEGFSAALITNGPLLSSTVGSYDWAG 376  
Db 301 NFELKTIQNLREKIFAIETGQTGSSSSSEHEMSQEGFSAALITNGPLLSSTVGSYDWAG 360  
Qy 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAALILNRVQSLVGLGAPRYCHIGLVAMFR 436  
Db 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAALILNRVQSLVGLGAPRYCHIGLVAMFR 420  
Qy 437 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQVSVCP 496  
Db 421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQVSVCP 480  
Qy 497 PRGQARWQCDVLYGEGQDPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
Db 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Qy 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDGLTMDGLVDTLVGAQGEVILLRSQ 616  
Db 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDGLTMDGLVDTLVGAQGEVILLRSQ 600  
Qy 617 PVLKVAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676  
Db 601 PVLKVAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
Qy 677 YDLALDSGRPHSRVAFNETKSTRTQVGLTQTCETLKLQLENPCIEDPVSIVLRNLF 736  
Db 661 YDLALDSGRPHSRVAFNETKSTRTQVGLTQTCETLKLQLENPCIEDPVSIVLRNLF 720  
Qy 737 SLVGTPLSAFNGNRPVLAEDAQRLLFTALFPPEKNCNDNI CQDDLSTTFPSMCLCLVWG 796  
Db 721 SLVGTPLSAFNGNRPVLAEDAQRLLFTALFPPEKNCNDNI CQDDLSTTFPSMCLCLVWG 780  
Qy 797 GPREFNVTVVRNDEGDSYRTQVTFPEPLDLSYKVSSTLQNSORSWRACSSASTEV 856  
Db 781 GPREFNVTVVRNDEGDSYRTQVTFPEPLDLSYKVSSTLQNSORSWRACSSASTEV 840  
Qy 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTPE 916  
Db 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTPE 900  
Qy 917 QLELPVYAVYVMVTSYGVSTKYLNTASNTSRVMOHGYOVNLCORSIPISLVLVPV 976  
Db 901 QLELPVYAVYVMVTSYGVSTKYLNTASNTSRVMOHGYOVNLCORSIPISLVLVPV 960  
Qy 977 RLNGTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1036  
Db 961 RLNGTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020  
Qy 1037 FFGIQEBFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDVSUPTLLPGQAFVRSQTET 1096  
Db 1021 FFGIQEBFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDVSUPTLLPGQAFVRSQTET 1080  
Qy 1097 KVEPPEVPNPLPLIVGSSVGLLALLIITAALYKLGFFKQYKXDMWSEGGPPGAEPQ 1153  
Db 1081 KVEPPEVPNPLPLIVGSSVGLLALLIITAALYKLGFFKQYKXDMWSEGGPPGAEPQ 1137

RESULT 10  
US-09-902-481A-3  
; Sequence 3, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
CURRENT APPLICATION NUMBER: US/09/902,481A  
CURRENT FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 60/216,600  
PRIOR FILING DATE: 2000-07-07  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 1137  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: synthetic  
US-09-902-481A-3

Query Match 98.0%; Score 5839; DB 10; Length 1137;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
Qy 17 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLVQCDYSTSCBEI 76  
Db 1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLVQCDYSTSCBEI 60  
Qy 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVYKGLCFGLFGSNLRQPOPK 136  
Db 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVYKGLCFGLFGSNLRQPOPK 120  
Qy 137 PPEARLGCQEDSDIAFLIDGSGSIIPHDPRMKGFVSTVMBOLKSKTLPISLMQVSEEF 196  
Db 121 PPEARLGCQEDSDIAFLIDGSGSIIPHDPRMKGFVSTVMBOLKSKTLPISLMQVSEEF 180  
Qy 197 RIHFTFKEFQNNPNRSLVKBPIQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 256  
Db 181 RIHFTFKEFQNNPNRSLVKBPIQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240  
Qy 257 TDGKFGDGLYEDVPEADREGVIRVVGIGDAFSEKSRQELNTIASRPHDHFQVN 316  
Db 241 TDGKFGDGLYEDVPEADREGVIRVVGIGDAFSEKSRQELNTIASRPHDHFQVN 300  
Qy 317 NFELKTIQNLREKIFAIETGQTGSSSSSEHEMSQEGFSAALITNGPLLSSTVGSYDWAG 376  
Db 301 NFELKTIQNLREKIFAIETGQTGSSSSSEHEMSQEGFSAALITNGPLLSSTVGSYDWAG 360  
Qy 377 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAALILNRVQSLVGLGAPRYCHIGLVAMFR 436  
Db 361 GVFLYTSKESKSTFNNTRVDSMDNDAYLGAAALILNRVQSLVGLGAPRYCHIGLVAMFR 420  
Qy 437 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQVSVCP 496  
Db 421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQVSVCP 480  
Qy 497 PRGQARWQCDVLYGEGQDPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
Db 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Qy 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDGLTMDGLVDTLVGAQGEVILLRSQ 616  
Db 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDGLTMDGLVDTLVGAQGEVILLRSQ 600  
Qy 617 PVLKVAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676  
Db 601 PVLKVAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
Qy 677 YDLALDSGRPHSRVAFNETKSTRTQVGLTQTCETLKLQLENPCIEDPVSIVLRNLF 736  
Db 661 YDLALDSGRPHSRVAFNETKSTRTQVGLTQTCETLKLQLENPCIEDPVSIVLRNLF 720  
Qy 737 SLVGTPLSAFNGNRPVLAEDAQRLLFTALFPPEKNCNDNI CQDDLSTTFPSMCLCLVWG 796  
Db 721 SLVGTPLSAFNGNRPVLAEDAQRLLFTALFPPEKNCNDNI CQDDLSTTFPSMCLCLVWG 780  
Qy 797 GPREFNVTVVRNDEGDSYRTQVTFPEPLDLSYKVSSTLQNSORSWRACSSASTEV 856  
Db 781 GPREFNVTVVRNDEGDSYRTQVTFPEPLDLSYKVSSTLQNSORSWRACSSASTEV 840

Db 781 GPREFNVTVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWELACESASSTEV 840  
QY 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTEP 916  
Db 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTEP 900  
QY 917 QLELPVKAVVMTVSHGVSTKYLNFNTASENTSRVMOHQYOVSNLQSRSLPISLVLFPV 976  
Db 901 QLELPVKAVVMTVSHGVSTKYLNFNTASENTSRVMOHQYOVSNLQSRSLPISLVLFPV 960  
QY 977 RLNQTVIWDROPVTFSENLSSTCHTKERLPKSHDFLABELKAPVNVCSIAVCRIQCIP 1036  
Db 961 RLNQTVIWDROPVTFSENLSSTCHTKERLPKSHDFLABELKAPVNVCSIAVCRIQCIP 1020  
QY 1037 PFGIOSEFNATLGNLSDWYIKTSHNLLIYSTAEILLPNDVSFTLLPQGFVRSQET 1096  
Db 1021 PFGIOSEFNATLGNLSDWYIKTSHNLLIYSTAEILLPNDVSFTLLPQGFVRSQET 1080  
QY 1097 KYEPFEVNPFLIYGVSSVGGLLALLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1153  
Db 1081 KYEPFEVNPFLIYGVSSVGGLLALLALITAAALYKLGFFKRYQKDMWSEGGPPGABPQ 1137

RESULT 11  
US-10-116-275-204  
; Sequence 204, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: 51067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 204  
; LENGTH: 1163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-275-204

Query Match 58.8%; Score 3503; DB 15; Length 1163;  
Best Local Similarity 61.3%; Pred. No. 1.1e-312;  
Matches 700; Conservative 138; Mismatches 297; Indels 6; Gaps 4;  
QY 5 VLLLTALTICHGPNLDTENAMTFOENARGFGQSVVQLQSRVWVGAPQEIIVANQSGSLY 64  
Db 8 LLLFTALATSLGFLNLTTELTAFRVDASGFSVWQYANSWVWVGAPQKITAANQTGGLY 67  
QY 65 QCDYSGSCPELRLOVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGLCF 124  
Db 68 QCGYSTGACEPIGLQVPPAVNMSLGLSLATTSPPOLLACGPTVHHECGRNMYLTGLCF 127  
QY 125 LFGSNLRQQPKFPEALRCQEDSDIAELIDCGSGIIPHDFRMEKFVSTVWOLKSKK 184  
Db 128 LLGPT--QLTQLPVRQBCPQEQDVELIDGSGSISSNFATMNFVRAVISQFQRES 185  
QY 185 TLPSLMQYSEEPRIHTFEKFNPNRSLVKPITOLLGRTHATGIRKVRVRELFTNIG 244  
Db 186 TQPSLMQFNSKPTHTTFEPRTSNPLSLASVHOLQFTYTATAIQNVHRLPHASVG 245  
QY 245 ARKNAFKILVITDGBKDPGLGYEDVIEADREGVIRYVIGVDAPRSEKROELNTIA 304  
Db 246 ASRDATKILLIVITDGBKESDLYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKLNDIA 305  
QY 305 SKPDRPHVQVNNFEALKTIONOLREKIATIEGTQSGSSSEHMSQEGFAALITNGP 364

Db 306 SKPSQEHIPKVEDFDALDKDIONQLKEKI PAI EGTETTTSSSSFELEMAQEGFSAVFTDPG 365  
QY 365 LLSTVGSYDWAGGVFLITSKEKSTPIMNTRVDSMDNDAYLGYAAAIILRRVQSLVAGP 424  
Db 366 VLGAAGSFTSGGAFLYPPNNSPTFIMNSQENVMRDSYLGYSTELAKWGVQSLVAGP 425  
QY 425 RYOHIGLVANFRONTGWESNANVKGFOICAYFGASLCSVDVDSNGSTDLVLICAPHYIE 484  
Db 426 RYQHTGAVFTQVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGGSTDLVLICAPHYIE 485  
QY 485 QTRGGQVSVCLPRGQARWQCDALVIGEQQPAGRFGAALTVLGDVNGDKLTDVAIGAP 544  
Db 486 QTRGGQVSVCLPRGQARWQCDALVIGEQQPAGRFGAALTVLGDVNGDKLTDVAIGAP 544  
QY 545 GEENRGAIVLPHGTSGSISPSHSQRIAGSLSPRLQYFQCSLSCGQDLTMDGLVLT 604  
Db 545 GEENRGAIVLPHGTSGSISPSHSQRIAGSLSPRLQYFQCSLSCGQDLTMDGLVLT 604  
QY 605 GAQGHVLLRSQPVLRVKAIMEFNPREVANVPCNDQVVKGEAGEVRVCLVQKSTRD 664  
Db 605 GARGVLLLRTRPVLWVGSMQFIPABIPRSAFECEQVWSEQLVQSNICLYIDKESKN 664  
QY 665 ELREGIOQSVVYDLALDSGRPHSRVFNFTNSTRTQTOVLGTLTQTCETLKLQPCIE 724  
Db 665 LLGSRDIQSSVTLDLALDPGLSPRATFQTKWRSLSRVRLGLKAHCENFNLLPSCVE 724  
QY 725 DPVSPVLRNLSVGTPLSAFAGNLRPVLADQRLFTALPPFKKNCNDNICDDLSIT 784  
Db 725 DSVPTITRLNFTLVGKPLLAFLNLRPMLAALAQRYFTASLPFEKNCADHICQDNLGIS 784  
QY 785 FSEMSLQVAVGGPREFNVTVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRW 844  
Db 785 FSPFGLKSLLVGSNLELNAEVMVWNGEDSYGTITITSHPAGLSYRYVREGQKQQLRSL 844  
QY 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVT 904  
Db 845 HLTCDASAPVG--SQSTWSTSCRIINHLIFRGAQITFLATPDVSPKAVLGDRLILTANVSS 902  
QY 905 ENNMPRTNTEFQLELPVKIYVMTVSHGVSTKYLNFNTAS--ENTSRVMOHQYOVSNLQ 963  
Db 903 ENNMPRTNTEFQLELPVKIYVMTVSHGVSTKYLNFNTAS--ENTSRVMOHQYOVSNLQ 962  
QY 964 RSLPISLVLPVRLNQVIMDRPQVTFSEKLSSTCHTKERLPKSHDFLABELKAPVNV 1023  
Db 963 RDLPVSNFVLPVNLQEAVMVDVSHPPQNPSPRLCSSEKTIAPPASDFLAHQNFVLD 1022  
QY 1024 SIACVQRIQCIPFPGIOSEFNATLGNLSDWYIKTSHNLLIYSTAEILLPNDVSFTLL 1083  
Db 1023 SIACVQRIQCIPFPGIOSEFNATLGNLSDWYIKTSHNLLIYSTAEILLPNDVSFTLL 1082  
QY 1084 PQGFVRSQETETKVEPPEVNPPLIYGVSSVGGLLALLALITAAALYKLGFFKRYQKDM 1143  
Db 1083 PQGFVRSQETETKVEPPEVNPPLIYGVSSVGGLLALLALITAAALYKLGFFKRYQKDM 1142  
QY 1144 E 1144  
Db 1143 E 1143

RESULT 12  
US-09-350-259-4  
; Sequence 4, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20020062008A1 Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497





QY 725 DVSPIVLRNLFSLVGTPLSAFAGNLRPLVLAEDAQRLEFALFPPEKNCNDNI CDDLSIT 784  
DB 722 DVVSPILHLNLFSLVRSPIPSQMLRPVLA VGSQDLFTASLPPEKNCQDGLCEGDLGVT 781  
QY 785 FSPMSLCLVYGGPREFNVTYVNDGEDSYRTOVTFPPPLDLSYKXVSTLQNRQSRW 844  
DB 782 LSPFGLQTLTVGSSLELNVIYVMNAGEDSYGTVWSLYIPAGLSHRVSGAQOPHQSAL 841  
QY 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904  
DB 842 RLACETV-PTDEG-LSSRCSVNHPHFEGSNGTFTVTFDVSYKATLGRMLRASASS 899  
QY 905 ENNMPRTNKTPEQLELPVKYAVMVTSHGVSITKYNP-TASENTSRVMQHOVQSNLQ 963  
DB 900 ENNKASSSKATFQLELPVKYAVYTMISROEBSTKYFNFAYSDEKMKAEHRVNNLSQ 959  
QY 964 RSLPISLVFLVPRLNQTVDWRPQVTFSENLSTCHTKRLPSHSDFLAELKAPVNC 1023  
DB 960 RDLAISINFVPLVLLNGVAVDVMWVMEAPSQL--PCVSEKPKQSDFLQISRPMLDC 1017  
QY 1024 SIACQRIQCDIPFGIOBEFNATLKNLSFDWIKTSHNHLIVSTABILFNDSVFTLL 1083  
DB 1018 SIADCLQFRCDVPSFVQSELDFTLKNLSFGVRETLQKKVLVSVVAEITFTDTSVYSL 1077  
QY 1084 PQCAFVRSOTETKVEPPEVDNPLPLIVGSSVGLLILLALITAAVYKLGFPKQYKXDMMS 1143  
DB 1078 PQCAFVRAQMEVLEEDVYNAIPIINGSSVGLLILLALITATLYKLGFPKHYKEMLE 1137  
QY 1144 E 1144  
DB 1138 D 1138

## RESULT 15

US-09-891-943-2  
; Sequence 2, Application US/09891943  
; Publication No. US20030077278A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; TITLE OF INVENTION: No. US20030077278A1 Human 2  
; FILE REFERENCE: 27666/35004  
; CURRENT APPLICATION NUMBER: US/09/891,943  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1161  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-891-943-2

Query Match 58.0%; Score 3455; DB 10; Length 1161;  
Best Local Similarity 59.9%; Pseq No. 2.9e-308;  
Matches 684; Conservative 162; Mismatches 287; Indels 8; Gaps 6;  
QY 5 VLLTALTLCHGFNLDENAMTFQENARGFQSVVQLQGRVYVVGAPQIEIVANQSGSLY 64  
DB 5 VLLLSVLASVHGFLNLDVEEPTIFQEDAGGFCQSVVQFGSRLVVGAPLEVVAAVNTQGRLY 64  
QY 65 QCDVSTGSCPEIRLOVPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVVKGLCF 124  
DB 65 DCAATGACQPIPLHRIPEAVNMSLGLTLAATNGSRLACGPTLHRCVGENSYSGKSL 124

QY 125 LFGSNLRQOPQKPEALRGCPQEDSDIAPLIDGSGSIIPHDPRMKBFVSTWELKKSK 184  
DB 125 LLGSRW-ELIQVTPDAPTECPHQENDIVPLIDGSGSIDQDNFQMKGFVQAVMGFEGD 183  
QY 185 TLPSLMQVSEEPRIHFTFKFQNNPNRSLVKPIITQLLQRTHTATCIRKVVARELFNITNG 244  
DB 184 TLPALMQVSNLKLHFTTQFATSPSQSLVDPIVQLKGLTTATGILLTVVQLFHHKG 243  
QY 245 ARNAPKILVITDEKFGDPLGYBDVTPEDRDEGVIRVIGVDAPREKRSQELNITIA 304  
DB 244 ARKSARKILVITDQKQVDPLEYSVDPVQAEGAGIIRYAIQVGHAFQPTARQELNITIS 303  
QY 305 SPPEDHVPQVNNFALKTIONLREKIFAIGTGTGSSSSPEHEMSQGFSAATNSGP 364  
DB 304 SAPPOHVPKYNFAALGSIQOLQEKIYAVSGTQSRASSSQHEMSQGFSTALTMDEL 363  
QY 365 LLSVGSVDWAGGVFLYTSKEKSTFINTRVDSMDNDAVLGYAAAIIILNRVQSLVIGAP 424  
DB 364 FLGAVGSPSWGGAFLYPPNMSPTFINNSQENVMDRDSYLGYSTELALWKGQNLVIGAP 423  
QY 425 RYCHICLVAMPQNTQWESNANVGTQIAYFGASLCSVDVDSNGSTDLVLIGAPHYTE 484  
DB 424 RYCHTQKAVITQVSRQWRKKAEBVTGTQIGSPFGLASLCSVDVDSGDTLLILIGAPHYTE 483  
QY 485 QTRGGQVSVCPFRGORARWQCDAVLYGEGQGPWCRFGAALTVLGDVNGDKLTDVAIGAP 544  
DB 484 QTRGGQVSVCPFRGORVQWQCDAVLRQBQGHWPGRFGAALTVLGDVNDKXLDVAIGAP 543  
QY 545 GEDNRGAVLPHGTSGSGISPSHSQRIAGSKLSRLOVFGQSLSGQDLTMDGLVDLTV 604  
DB 544 GQENRGAVLPHGASBSGISPSHSQRIASSQLSRLOVFGQALSGQDLTQDGLMDLAV 603  
QY 605 GAQGHVLLLRSPQVLRVKAIMEFNPREVARNYFECNDQVVGKEAGEVRLVHVGKSTRD 664  
DB 604 GARGQVLLLRSLPVLKGVAMRFSPVVEKAVYVRCWEKPSALEAGDATVCLTIQKSLD 663  
QY 665 RLREGIOQVWYVDLALDSGRPHSAVNETKNSRTOQVGLGTCTCEKTLQLPNCLE 724  
DB 664 QL--GDIOQSVRFDLALDQRLTSAIFNETKNPTLTRKTLGLGTHCSTKLLLEDCVE 721  
QY 725 DVSPIVLRNLFSLVGTPLSAFAGNLRPLVLAEDAQRLEFALFPPEKNCNDNICDDLSIT 784  
DB 722 DVVSPILHLNLFSLVRSPIPSQMLRPVLA VGSQDLFTASLPPEKNCQDGLCEGDLGVT 781  
QY 785 FSPMSLCLVYGGPREFNVTYVNDGEDSYRTOVTFPPPLDLSYKXVSTLQNRQSRW 844  
DB 782 LSPFGLQTLTVGSSLELNVIYVMNAGEDSYGTVWSLYIPAGLSHRVSGAQOPHQSAL 841  
QY 845 RLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTS 904  
DB 842 RLACETV-PTDEG-LSSRCSVNHPHFEGSNGTFTVTFDVSYKATLGRMLRASASS 899  
QY 905 ENNMPRTNKTPEQLELPVKYAVMVTSHGVSITKYNP-TASENTSRVMQHOVQSNLQ 963  
DB 900 ENNKASSSKATFQLELPVKYAVYTMISROEBSTKYFNFAYSDEKMKAEHRVNNLSQ 959  
QY 964 RSLPISLVFLVPRLNQTVDWRPQVTFSENLSTCHTKRLPSHSDFLAELKAPVNC 1023  
DB 960 RDLAISINFVPLVLLNGVAVDVMWVMEAPSQL--PCVSEKPKQSDFLQISRPMLDC 1017  
QY 1024 SIACQRIQCDIPFGIOBEFNATLKNLSFDWIKTSHNHLIVSTABILFNDSVFTLL 1083  
DB 1018 SIADCLQFRCDVPSFVQSELDFTLKNLSFGVRETLQKKVLVSVVAEITFTDTSVYSL 1077  
QY 1084 PQCAFVRSOTETKVEPPEVDNPLPLIVGSSVGLLILLALITAAVYKLGFPKQYKXDMMS 1143  
DB 1078 PQCAFVRAQMEVLEEDVYNAIPIINGSSVGLLILLALITATLYKLGFPKHYKEMLE 1137  
QY 1144 E 1144  
DB 1138 D 1138

us-09-902-481b-1.rapb

Tue Jun 8 05:59:06 2004

Search completed: June 7, 2004, 17:38:40  
Job time : 43.4378 secs



QY 61 GSLYCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120  
DB 61 GSLYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120  
QY 121 GLCFLFGSNLRQOQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFPVSTVMEQL 180  
DB 121 GLCFLFGSNLRQOQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFPVSTVMEQL 180  
QY 181 KKSRTLPSLMQYSEEFRIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFN 240  
DB 181 KKSRTLPSLMQYSEEFRIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFN 240  
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQEL 300  
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQEL 300  
QY 301 NTIASKPRDRHVQVNNPFAELKTIONLREKIPALEGTOTGSSSSFEHMSQEGFSAIT 360  
DB 301 NTIASKPRDRHVQVNNPFAELKTIONLREKIPALEGTOTGSSSSFEHMSQEGFSAIT 360  
QY 361 SNGFLLSTVGSYDWDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLV 420  
DB 361 SNGFLLSTVGSYDWDAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLV 420  
QY 421 LGAPRYQHIGLVANFRONTGMESNNVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAP 480  
DB 421 LGAPRYQHIGLVANFRONTGMESNNVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAP 480  
QY 481 HYEQTRGGQVSVCPFLRGQARQCDVAVLYGEOQPGWRFGAALTIVLGDVNGDKLTDVA 540  
DB 481 HYEQTRGGQVSVCPFLRGQARQCDVAVLYGEOQPGWRFGAALTIVLGDVNGDKLTDVA 540  
QY 541 IGAPGEDNRGANVLYHGTSGSGISPSHSORIASGKLSPLRYPGQSLSGQDLTMDGLV 600  
DB 541 IGAPGEDNRGANVLYHGTSGSGISPSHSORIASGKLSPLRYPGQSLSGQDLTMDGLV 600  
QY 601 DLTGACQHVLLLRSPQVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVVRVCLHVQK 660  
DB 601 DLTGACQHVLLLRSPQVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVVRVCLHVQK 660  
QY 720 STDRRLREGQIQSVVTVYDLALDSGRPHSAVFNETKNSRRQTVGLGLTQCTETKLQLP 720  
DB 720 STDRRLREGQIQSVVTVYDLALDSGRPHSAVFNETKNSRRQTVGLGLTQCTETKLQLP 720  
QY 721 NCIEDPVSPIVLRNLSVGTPLSAFGLNLPVLAEDAQRFTALPPEKNCGNNDICQDD 780  
DB 721 NCIEDPVSPIVLRNLSVGTPLSAFGLNLPVLAEDAQRFTALPPEKNCGNNDICQDD 780  
QY 781 LSITFSFMSLDCLVVGCPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR 840  
DB 781 LSITFSFMSLDCLVVGCPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNR 840  
QY 841 QRSRLACESASSTEVSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKA 900  
DB 841 QRSRLACESASSTEVSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKA 900  
QY 901 NVTSENMPRTNTEFQLELPVKYAVTVMTVSHGVSTKYLNFTASENTSRVMQHQVSN 960  
DB 901 NVTSENMPRTNTEFQLELPVKYAVTVMTVSHGVSTKYLNFTASENTSRVMQHQVSN 960  
QY 961 LGQBSLPSILVFLVPLNQTIVDRQVTFSENLSSTCTKRLPSHSDPFLAELKAPV 1020  
DB 961 LGQBSLPSILVFLVPLNQTIVDRQVTFSENLSSTCTKRLPSHSDPFLAELKAPV 1020  
QY 1021 VNCISIAVCQRIQCDIPFFGIEQEFNATLKGNLSPDWYIKTSHNLLITVSTAEILFNDVSF 1080  
DB 1021 VNCISIAVCQRIQCDIPFFGIEQEFNATLKGNLSPDWYIKTSHNLLITVSTAEILFNDVSF 1080  
QY 1081 TLLPQGAQVRSQETKTVSEFPVNPFLPLIVGSSVGLLALLIATLYKLGFFKQYKD 1140  
DB 1081 TLLPQGAQVRSQETKTVSEFPVNPFLPLIVGSSVGLLALLIATLYKLGFFKQYKD 1140

QY 1141 MMSEGGPPGRNEPQ 1153  
DB 1141 MMSEGGPPGRNEPQ 1153

## RESULT 2

US-08-286-889-3

; Sequence 3, Application US/08286889

; Patent No. 5470953

; GENERAL INFORMATION:

; APPLICANT: Gallatin, W. Mich

; APPLICANT: Van der Vieren, Monica

; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &amp; Borun

; STREET: 233 South Wacker Drive, 6300 Sear Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/286,889

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/173,497

; FILING DATE: 23-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: P38,659

; REFERENCE/DOCKET NUMBER: 27866/32168

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-474-6300

; TELEFAX: 312-474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1153 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-286-889-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIVAANOR 60

DB 1 MALRVLLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIVAANOR 60

QY 61 GSLYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120

DB 61 GSLYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120

QY 121 GLCFLFGSNLRQOQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFPVSTVMEQL 180

DB 121 GLCFLFGSNLRQOQKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFPVSTVMEQL 180

QY 181 KKSRTLPSLMQYSEEFRIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFN 240

DB 181 KKSRTLPSLMQYSEEFRIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGIRKVVRELFN 240

QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQEL 300

DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQEL 300

301 NTIASKPRDRHVQVNNFEALKTIONQREKIPAIETGTGTGSSSSFEHMSQSGFSAAIT 360  
301 NTIASKPRDRHVQVNNFEALKTIONQREKIPAIETGTGTGSSSSFEHMSQSGFSAAIT 360  
361 SNGPLLSVGSYDWAGGVFLYTSKSTFTINTRVDSMDNDAYLGVAAAIILNRVQSIV 420  
361 SNGPLLSVGSYDWAGGVFLYTSKSTFTINTRVDSMDNDAYLGVAAAIILNRVQSIV 420  
421 LGAPRYQHIGLVAMFRQNTGMSNANVKGTOIGAVFGASLCSVDVDSNGSTDVLIGAP 480  
421 LGAPRYQHIGLVAMFRQNTGMSNANVKGTOIGAVFGASLCSVDVDSNGSTDVLIGAP 480  
481 HYETQTRGGQVSVCLPRGQARWQCDVLYGBOGQWGRFGAALTVLGDVNGDKLTDVA 540  
481 HYETQTRGGQVSVCLPRGQARWQCDVLYGBOGQWGRFGAALTVLGDVNGDKLTDVA 540  
541 IGAPGEDNRGAYLVHGTGSGSISHSORIASKLSRLQVFGOSLSGGQDLTWGGLV 600  
541 IGAPGEDNRGAYLVHGTGSGSISHSORIASKLSRLQVFGOSLSGGQDLTWGGLV 600  
601 DLTGVAQGHVLLRSQPLRVKKAIMEFNPFEVARNVPECNDQVVGKEAGEVRVCLHVQK 660  
601 DLTGVAQGHVLLRSQPLRVKKAIMEFNPFEVARNVPECNDQVVGKEAGEVRVCLHVQK 660  
661 STDRLRREGQIQSVVYTDIALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQLP 720  
661 STDRLRREGQIQSVVYTDIALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQLP 720  
721 NCIEDPVSPIVLRLNLSVLTGTPLSAQNLRPVLAEDAQLFTALFPFKXNCGNDNICQDD 780  
721 NCIEDPVSPIVLRLNLSVLTGTPLSAQNLRPVLAEDAQLFTALFPFKXNCGNDNICQDD 780  
781 LSITPFSMLDCLVWGPREFNVTVVRNDGDSYRTQVTPFPPLDLSYRKVSTLQNS 840  
781 LSITPFSMLDCLVWGPREFNVTVVRNDGDSYRTQVTPFPPLDLSYRKVSTLQNS 840  
841 ORSWRLACSSASSTVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA 900  
841 ORSWRLACSSASSTVSGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKA 900  
901 NVTSENMPRTNKTETQLELPVKYAVVMTVSHGVSTKYLNFTASNTSRVMOHQVQSN 960  
901 NVTSENMPRTNKTETQLELPVKYAVVMTVSHGVSTKYLNFTASNTSRVMOHQVQSN 960  
961 LGORSIPISLVFLVPLVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020  
961 LGORSIPISLVFLVPLVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPV 1020  
1021 VNCISIAVCORIQCIDIFFGTOEENATLKGNTLSFDWYIKTSHNHLIIVSTABILFNDVSV 1080  
1021 VNCISIAVCORIQCIDIFFGTOEENATLKGNTLSFDWYIKTSHNHLIIVSTABILFNDVSV 1080  
1081 TLLPGQGAFFRSQTEKVPPEVNDPLPLVGSVGGILLALITAALYKLGFKKQYKD 1140  
1081 TLLPGQGAFFRSQTEKVPPEVNDPLPLVGSVGGILLALITAALYKLGFKKQYKD 1140  
1141 MNSEGGPPGAEPQ 1153  
1141 MNSEGGPPGAEPQ 1153

RESULT 3  
US-08-485-618-3  
Sequence 3, Application US/08485618  
Patent No. 5728533  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,618  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32797  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-618-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTLCGPNLDTENAMTFOENARGFQSVVVLQGSRRVVGAPQEIIVAANQR 60  
DB 1 MALRVLLLTALTLCGPNLDTENAMTFOENARGFQSVVVLQGSRRVVGAPQEIIVAANQR 60  
QY 61 GSLYQCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120  
DB 61 GSLYQCDYSTGSCPEIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVK 120  
QY 121 GLCFLFGSNLRQOPQKPFPEARLRCQPEDSDIAFLIDGSGSIIPHPDFRKNKEFVSTMBSQL 180  
DB 121 GLCFLFGSNLRQOPQKPFPEARLRCQPEDSDIAFLIDGSGSIIPHPDFRKNKEFVSTMBSQL 180  
QY 181 KKSKTFLSLMOYSEBEFRIFHTFKCFQNNPNRSLVKPITQLLGRTHHTATGIRKVVRELFN 240  
DB 181 KKSKTFLSLMOYSEBEFRIFHTFKCFQNNPNRSLVKPITQLLGRTHHTATGIRKVVRELFN 240  
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQEL 300  
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQEL 300  
QY 301 NTIASKPRDRHVQVNNFEALKTIONQREKIPAIETGTGTGSSSSFEHMSQSGFSAAIT 360  
DB 301 NTIASKPRDRHVQVNNFEALKTIONQREKIPAIETGTGTGSSSSFEHMSQSGFSAAIT 360  
QY 361 SNGPLLSVGSYDWAGGVFLYTSKSTFTINTRVDSMDNDAYLGVAAAIILNRVQSIV 420  
DB 361 SNGPLLSVGSYDWAGGVFLYTSKSTFTINTRVDSMDNDAYLGVAAAIILNRVQSIV 420  
QY 421 LGAPRYQHIGLVAMFRQNTGMSNANVKGTOIGAVFGASLCSVDVDSNGSTDVLIGAP 480

421 LGAPRYOHIGLVANFRONTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480  
481 HYETQTRGGQSVCPPLRQORARWQCDVLYGEGQPGWGFAGALTIVLGDVNGDKLTDVA 540  
481 HYETQTRGGQSVCPPLRQORARWQCDVLYGEGQPGWGFAGALTIVLGDVNGDKLTDVA 540  
541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600  
541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600  
601 DLTGAGQSHVLLRSQVLRKALMEFNPREVARNPFCNDQVVKGEAGEVRVCLHVQK 660  
601 DLTGAGQSHVLLRSQVLRKALMEFNPREVARNPFCNDQVVKGEAGEVRVCLHVQK 660  
661 STRDLREGQIQSVVYDLDLDSRPHSRVAFNETKNSSTRQTOVLGLTQTCETLKQLP 720  
661 STRDLREGQIQSVVYDLDLDSRPHSRVAFNETKNSSTRQTOVLGLTQTCETLKQLP 720  
721 NCIEDPVPVLRNPLNPSLGTPLSAFGNLRPVLAEDAQRLLFTALFPPEKNCMDNICQDD 780  
721 NCIEDPVPVLRNPLNPSLGTPLSAFGNLRPVLAEDAQRLLFTALFPPEKNCMDNICQDD 780  
781 LSITPFSMSLDCLVVGPREFNTVTVRNDEGDSYRTQVTFEPLDLSYKXVSTLQNORS 840  
781 LSITPFSMSLDCLVVGPREFNTVTVRNDEGDSYRTQVTFEPLDLSYKXVSTLQNORS 840  
841 QRSWRLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKA 900  
841 QRSWRLACESASSTEVSGALKSTSCSINHPIPPENSEVTFNITPDVDSKASLGNKLLKA 900  
901 NVTSENNMPNKTKEFOLELPVKYAVVMVTSYKLNFTASENTSRVMQHOYOVSN 960  
901 NVTSENNMPNKTKEFOLELPVKYAVVMVTSYKLNFTASENTSRVMQHOYOVSN 960  
961 LGORSPLISLVLVPLVRLNQTIVINDRPQVTFPSENLSSTCHTKERLPSHSDFLAELKAPV 1020  
961 LGORSPLISLVLVPLVRLNQTIVINDRPQVTFPSENLSSTCHTKERLPSHSDFLAELKAPV 1020  
1021 VNCISAVCORIQDIPFGIOEBFNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSF 1080  
1021 VNCISAVCORIQDIPFGIOEBFNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSF 1080  
1081 TLLPQCGAFVRSQETKVEPPEVNPPLPIVGGSVGGLLIALLITAALYKLGFFKQYKD 1140  
1081 TLLPQCGAFVRSQETKVEPPEVNPPLPIVGGSVGGLLIALLITAALYKLGFFKQYKD 1140  
1141 MWSEGGPPGAPQ 1153  
1141 MWSEGGPPGAPQ 1153

RESULT 4  
US-08-362-652-3  
; Sequence 3, Application US/08362652  
; Patent No. 5766850  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/362,652  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32391  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-362-652-3

Query Match 100.0%; Score 5956; DB 1; Length 1153;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MALRVLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR 60  
Db 1 MALRVLLTALTLCGHNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR 60  
Qy 61 GSLYQCDYSTGCEPRLQVPEAVNMSLGLSLAATSPOLLACQPTVHQCSENYTK 120  
Db 61 GSLYQCDYSTGCEPRLQVPEAVNMSLGLSLAATSPOLLACQPTVHQCSENYTK 120  
Qy 121 GUCFLPGSNLRQOPKPFPEALRGCCQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQL 180  
Db 121 GUCFLPGSNLRQOPKPFPEALRGCCQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQL 180  
Qy 181 KSKTILFSLMXYSEFRIFHTFKFQNNPNPSLVKPIITQLLGRTHATGIRKRVRELFN 240  
Db 181 KSKTILFSLMXYSEFRIFHTFKFQNNPNPSLVKPIITQLLGRTHATGIRKRVRELFN 240  
Qy 241 ITNGARKNAFKILVITDGEKEDPLGYEDVLPEDREGVIRVYVGVGDAFRSEKSRQEL 300  
Db 241 ITNGARKNAFKILVITDGEKEDPLGYEDVLPEDREGVIRVYVGVGDAFRSEKSRQEL 300  
Qy 301 NTIASKPPRDHVPQVNNFEALKTIONQLEKIFAIEGTQTGSSSFHEMSQEGFSAIT 360  
Db 301 NTIASKPPRDHVPQVNNFEALKTIONQLEKIFAIEGTQTGSSSFHEMSQEGFSAIT 360  
Qy 361 SNGPLSTVSGVDWAGGVLYTSKEKSTEINTRVDSMDNDAYLGYAAAILLRNVQSLV 420  
Db 361 SNGPLSTVSGVDWAGGVLYTSKEKSTEINTRVDSMDNDAYLGYAAAILLRNVQSLV 420  
Qy 421 LGAPRYOHIGLVANFRONTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480  
Db 421 LGAPRYOHIGLVANFRONTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAP 480  
Qy 481 HYETQTRGGQSVCPPLRQORARWQCDVLYGEGQPGWGFAGALTIVLGDVNGDKLTDVA 540  
Db 481 HYETQTRGGQSVCPPLRQORARWQCDVLYGEGQPGWGFAGALTIVLGDVNGDKLTDVA 540  
Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600  
Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLV 600  
Qy 601 DLTGAGQSHVLLRSQVLRKALMEFNPREVARNPFCNDQVVKGEAGEVRVCLHVQK 660  
Db 601 DLTGAGQSHVLLRSQVLRKALMEFNPREVARNPFCNDQVVKGEAGEVRVCLHVQK 660

601 DLTGAGQHVVLLRSQPVLRVKAIMEFNPRAVARNVFCNDQVVGKKEAGEVRVCLHVOK 660  
661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKSTRQVGLTQTCETLKLQLP 720  
661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKSTRQVGLTQTCETLKLQLP 720  
721 NCIEDVSPVILRLNLSVGLTSLASGNLRPVLAEDAQRFTALPFPEKNCNDNICQDD 780  
721 NCIEDVSPVILRLNLSVGLTSLASGNLRPVLAEDAQRFTALPFPEKNCNDNICQDD 780  
781 LSTTFMSLDCLVVGCPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNR 840  
781 LSTTFMSLDCLVVGCPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNR 840  
841 QRSWRLACESASSTVSGALKSTSCINHPDPENSEVFNITFDVDSKASIGKLLKA 900  
841 QRSWRLACESASSTVSGALKSTSCINHPDPENSEVFNITFDVDSKASIGKLLKA 900  
901 NVTSENMPNTKTEFOLELPVKYAVVMTVSHGVSTKYNLTASNTSRVMOHQVSN 960  
901 NVTSENMPNTKTEFOLELPVKYAVVMTVSHGVSTKYNLTASNTSRVMOHQVSN 960  
961 LQORSIPISVPLVPLVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPV 1020  
961 LQORSIPISVPLVPLVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKPV 1020  
1021 VNCIAVCQRIQCDIPFGIOEENFATLKNLSPDMVYIKTSHNHLIVSTABILFNDVVF 1080  
1021 VNCIAVCQRIQCDIPFGIOEENFATLKNLSPDMVYIKTSHNHLIVSTABILFNDVVF 1080  
1081 TLLPGQAFVRSQTEKVPFEPVNPPLIVGSSVGLLALLITAAKYKLGFFKQYKD 1140  
1081 TLLPGQAFVRSQTEKVPFEPVNPPLIVGSSVGLLALLITAAKYKLGFFKQYKD 1140  
1141 MMSEGGPPGAEPQ 1153  
1141 MMSEGGPPGAEPQ 1153

RESULT 5  
US-08-605-672-3  
Sequence 3, Application US/08605672  
Patent No. 5817515  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,672  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-605-672-3

Query Match 100.0%; Score 5956; DB 2; Length 1153;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLTALTLCCHGNLDTENAMTFQENARGFCQSVVQLQGSRVVVGAPQEIIVAANOR 60  
DB 1 MALRVLLTALTLCCHGNLDTENAMTFQENARGFCQSVVQLQGSRVVVGAPQEIIVAANOR 60  
QY 61 GSLYQCDYSTGSCCEPIRLQVPEAVNMSLGLSLAATTSPQALLACGPTVHQTCSNTYVK 120  
DB 61 GSLYQCDYSTGSCCEPIRLQVPEAVNMSLGLSLAATTSPQALLACGPTVHQTCSNTYVK 120  
QY 121 GLCFLPGNLRQOPKPFALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180  
DB 121 GLCFLPGNLRQOPKPFALRGCPQSDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQL 180  
QY 181 KSKTFLSLMQYSEBFRHFTFKBPQNNPNPSLVKPIITQLGRTHATATGIRKVVRELFN 240  
DB 181 KSKTFLSLMQYSEBFRHFTFKBPQNNPNPSLVKPIITQLGRTHATATGIRKVVRELFN 240  
QY 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVTPEDADREGVIRVYGVGDAPSEKSRQEL 300  
DB 241 ITNGARKNAFKILVITDGEKFGDPLGYEDVTPEDADREGVIRVYGVGDAPSEKSRQEL 300  
QY 301 NTIASKPPRDHVFNQNNFEALKTIQNLREKIFALTEGTGTGSSSSPHEHMSQEGFSAAT 360  
DB 301 NTIASKPPRDHVFNQNNFEALKTIQNLREKIFALTEGTGTGSSSSPHEHMSQEGFSAAT 360  
QY 361 SNGPLLSTVGSYDAGGVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAIILRNRVQSLV 420  
DB 361 SNGPLLSTVGSYDAGGVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAIILRNRVQSLV 420  
QY 421 LGAPRYQHIGLVAMPFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAP 480  
DB 421 LGAPRYQHIGLVAMPFRONTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAP 480  
QY 481 HYYEQTRGGQSVCPPLRGQARWQCDVLYGEOQCPWGRFGAALTVLGDVNGDKLTDVA 540  
DB 481 HYYEQTRGGQSVCPPLRGQARWQCDVLYGEOQCPWGRFGAALTVLGDVNGDKLTDVA 540  
QY 541 ICAPGEENRGAVYLPFGTSGSGISPSHSORITAGSKLSPRLQYFGQSLGGQDLTMDGLV 600  
DB 541 ICAPGEENRGAVYLPFGTSGSGISPSHSORITAGSKLSPRLQYFGQSLGGQDLTMDGLV 600  
QY 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPRAVARNVFCNDQVVGKKEAGEVRVCLHVOK 660  
DB 601 DLTVGAQGHVLLRSQPVLRVKAIMEFNPRAVARNVFCNDQVVGKKEAGEVRVCLHVOK 660  
QY 661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKSTRQVGLTQTCETLKLQLP 720  
DB 661 STRDLREGQIQSVVYDIALDSGRPHSRAVFNETKSTRQVGLTQTCETLKLQLP 720  
QY 721 NCIEDVSPVILRLNLSVGLTSLASGNLRPVLAEDAQRFTALPFPEKNCNDNICQDD 780  
DB 721 NCIEDVSPVILRLNLSVGLTSLASGNLRPVLAEDAQRFTALPFPEKNCNDNICQDD 780

Qy 781 LSITFSFMSLCLVVGPGREFNVTVVRNDGDSYRTQVTPFFPLDLSYRKVSTLQNR 840  
Db 781 LSITFSFMSLCLVVGPGREFNVTVVRNDGDSYRTQVTPFFPLDLSYRKVSTLQNR 840  
Qy 841 QRSWELACESASTVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900  
Db 841 QRSWELACESASTVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900  
Qy 901 NVTSENNMPTNKTEFFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVYMQHGYOVSN 960  
Db 901 NVTSENNMPTNKTEFFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVYMQHGYOVSN 960  
Qy 961 LQORSLPISLVLVRLNQVIMDRPOVTVSENLSSTCHTKERLPDSHDFLAEIRKAPV 1020  
Db 961 LQORSLPISLVLVRLNQVIMDRPOVTVSENLSSTCHTKERLPDSHDFLAEIRKAPV 1020  
Qy 1021 VNCSTAVCQRIQCDIPFGIOEFNATLKGNLSPDWYIKTSHNHLIIIVSTAELIFNDSVF 1080  
Db 1021 VNCSTAVCQRIQCDIPFGIOEFNATLKGNLSPDWYIKTSHNHLIIIVSTAELIFNDSVF 1080  
Qy 1081 TLLPQGAFAVSQTKVPEPEVNPPLFIVGSSVGGLLLLALITAAALYKLGPFKQVYD 1140  
Db 1081 TLLPQGAFAVSQTKVPEPEVNPPLFIVGSSVGGLLLLALITAAALYKLGPFKQVYD 1140  
Qy 1141 MWSEGGPPGABQ 1153  
Db 1141 MWSEGGPPGABQ 1153

RESULT 6  
US-08-482-293A-3  
Sequence 3, Application US/08482293A  
Patent No. 5831029  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Marehall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,293A  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-482-293A-3  
Query Match  
Best Local Similarity 100.0%; Score 5956; DB 2; Length 1153;  
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MALRVLLLTALTICCHFNLDENAMTQENARFGQSVVOLQGRVGVVGAPOQIVANQR 60  
Db 1 MALRVLLLTALTICCHFNLDENAMTQENARFGQSVVOLQGRVGVVGAPOQIVANQR 60  
Qy 61 GSIVQCDYSGSCPTLRLOVPEAVNMSGLSLAATTPPOLACGPTVHQTCESTYVK 120  
Db 61 GSIVQCDYSGSCPTLRLOVPEAVNMSGLSLAATTPPOLACGPTVHQTCESTYVK 120  
Qy 121 GLCFLFGSNLRQOPKQPEALRGCPQEDSDIAFLIDSGSII PHDFRMEKPFYSTVMEQL 180  
Db 121 GLCFLFGSNLRQOPKQPEALRGCPQEDSDIAFLIDSGSII PHDFRMEKPFYSTVMEQL 180  
Qy 181 KXSKTLFSLMOYSEEFRIHFTKPEFQNNPFRSLVKPIITOLLGRTHATTGIRKVVRELFN 240  
Db 181 KXSKTLFSLMOYSEEFRIHFTKPEFQNNPFRSLVKPIITOLLGRTHATTGIRKVVRELFN 240  
Qy 241 ITNGARKNAFKILVITDGEKPGDPLGYEDVIPEADREGVIRYVIGVGDAPRSEKSRQEL 300  
Db 241 ITNGARKNAFKILVITDGEKPGDPLGYEDVIPEADREGVIRYVIGVGDAPRSEKSRQEL 300  
Qy 301 NTIASKPPRDHVFQVNNFALKTIQNLBEKIFAIETGOTGSSSSSEHEMSQSGFSAIT 360  
Db 301 NTIASKPPRDHVFQVNNFALKTIQNLBEKIFAIETGOTGSSSSSEHEMSQSGFSAIT 360  
Qy 361 SNGPLSLTSGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV 420  
Db 361 SNGPLSLTSGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLV 420  
Qy 421 LGAPRYQHIGLVAMFRONTGWNESNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAP 480  
Db 421 LGAPRYQHIGLVAMFRONTGWNESNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAP 480  
Qy 481 HYTEQRTGGQVSVCPILPRGQARQWCDVLYGQGGQPGWGFGAALTIVLGVNDGKLTQV 540  
Db 481 HYTEQRTGGQVSVCPILPRGQARQWCDVLYGQGGQPGWGFGAALTIVLGVNDGKLTQV 540  
Qy 541 IGAPGEDNREGAVYLFHGTSGSGISPSHSORIIAGSKLSPRLQYFGQSLSGQDLTMDGLV 600  
Db 541 IGAPGEDNREGAVYLFHGTSGSGISPSHSORIIAGSKLSPRLQYFGQSLSGQDLTMDGLV 600  
Qy 601 DLTVGAQGHVLLRSQPLRVKKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQK 660  
Db 601 DLTVGAQGHVLLRSQPLRVKKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQK 660  
Qy 661 STRDLRREGQIQSVTVVLDALDSGRPHSRVAVPNETKNSRTRQVTLGLTQTCETLKLQLP 720  
Db 661 STRDLRREGQIQSVTVVLDALDSGRPHSRVAVPNETKNSRTRQVTLGLTQTCETLKLQLP 720  
Qy 721 NCIEDPVSPVILNFSLVCTPLSAFGNLEPVLAEQAQLFTALPPEKKGNDNICQDD 780  
Db 721 NCIEDPVSPVILNFSLVCTPLSAFGNLEPVLAEQAQLFTALPPEKKGNDNICQDD 780  
Qy 781 LSITFSFMSLCLVVGPGREFNVTVVRNDGDSYRTQVTPFFPLDLSYRKVSTLQNR 840  
Db 781 LSITFSFMSLCLVVGPGREFNVTVVRNDGDSYRTQVTPFFPLDLSYRKVSTLQNR 840  
Qy 841 QRSWELACESASTVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900  
Db 841 QRSWELACESASTVSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900  
Qy 901 NVTSENNMPTNKTEFFQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVYMQHGYOVSN 960

Db 901 NVTSENNMPTNKTEFQLELPVKAVYVWVTSHGVSSTKYLNFPTASNTSRVWQHOYQVSN 960  
 Qy 961 LGQSLPISLVFLVRLNQTVINDRPQVTFSENLSTCTKTERLPSSHSDFLAELRKAPV 1020  
 Db 961 LGQSLPISLVFLVRLNQTVINDRPQVTFSENLSTCTKTERLPSSHSDFLAELRKAPV 1020  
 Qy 1021 VNCISIAVCORIQCDIPFPGIOEEFNATLKGMLSDWYIKTSHNHLIVSTAEILFNDVSF 1080  
 Db 1021 VNCISIAVCORIQCDIPFPGIOEEFNATLKGMLSDWYIKTSHNHLIVSTAEILFNDVSF 1080  
 Qy 1081 TLLPQCGAFVRSQYETKVEPEVENPLIIVGSSVGGILLALITAAALYKLGPFKQRYKD 1140  
 Db 1081 TLLPQCGAFVRSQYETKVEPEVENPLIIVGSSVGGILLALITAAALYKLGPFKQRYKD 1140  
 Qy 1141 MMSEGGPPGASPO 1153  
 Db 1141 MMSEGGPPGASPO 1153

RESULT 7

; US-08-943-363-3  
 ; Sequence 3, Application US/08943363  
 ; Patent No. 5837478  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, W. Michael  
 ; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit  
 ; NUMBER OF SEQUENCES: 114  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 South Wacker Drive, 6300 Sear Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/943,363  
 ; FILING DATE:  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/173,497  
 ; FILING DATE: 23-DEC-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/286,889  
 ; FILING DATE: 5-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/362,652  
 ; FILING DATE: 21-DEC-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Williams Jr., Joseph A.  
 ; REGISTRATION NUMBER: 38,659  
 ; REFERENCE/DOCKET NUMBER: 27866/32584  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448  
 ; TELEX: 25-3856  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1153 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-943-363-3

Query Match 100.0%; Score 5956; DB 2; Length 1153;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARGFGQSVVQLQSGSRVVVGAPQBIIVANQR 60  
 Db 1 MALRVLLLTALTLCCHGFNLDTENAMTFQENARGFGQSVVQLQSGSRVVVGAPQBIIVANQR 60  
 Qy 61 GSIYQCDYSGSCERPLQVPEAVNMSGLSLAATTSPPQILLACGTPVHQTCSNTYVK 120  
 Db 61 GSIYQCDYSGSCERPLQVPEAVNMSGLSLAATTSPPQILLACGTPVHQTCSNTYVK 120  
 Qy 121 GLCFLFGSNLRQOPKPFPEALRCQPOEDSDIAFLIDSGSGSIIIPHDPRMKEFVSTVMEQL 180  
 Db 121 GLCFLFGSNLRQOPKPFPEALRCQPOEDSDIAFLIDSGSGSIIIPHDPRMKEFVSTVMEQL 180  
 Qy 181 KKSKTILFSLMOYSEEFRIHETFKSEFQNNPNRSLVKPITOLLGRTHTATGIRKVRLEFN 240  
 Db 181 KKSKTILFSLMOYSEEFRIHETFKSEFQNNPNRSLVKPITOLLGRTHTATGIRKVRLEFN 240  
 Qy 241 ITNGARKNAFKILVWITDGEKFGDPLGYEDVIFEADREGVIRVIVGDAFRSEKSRQEL 300  
 Db 241 ITNGARKNAFKILVWITDGEKFGDPLGYEDVIFEADREGVIRVIVGDAFRSEKSRQEL 300  
 Qy 301 NTIASKPPRDHVEQVANNPEALKTIQNLREKIFAIETGTQGTGSSSSFEHEHSGQFSAIT 360  
 Db 301 NTIASKPPRDHVEQVANNPEALKTIQNLREKIFAIETGTQGTGSSSSFEHEHSGQFSAIT 360  
 Qy 361 SNGPILLSVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLENRVQSLV 420  
 Db 361 SNGPILLSVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLENRVQSLV 420  
 Qy 421 LGAPRYCHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480  
 Db 421 LGAPRYCHIGLVAMFRQNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAP 480  
 Qy 481 HYETQTRGGQVSVCPPLRGORARWQCDVLYGSGOGPWGRFGAALTIVLGVNCGDKLTDVA 540  
 Db 481 HYETQTRGGQVSVCPPLRGORARWQCDVLYGSGOGPWGRFGAALTIVLGVNCGDKLTDVA 540  
 Qy 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSCGQDLTMDGLV 600  
 Db 541 IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSCGQDLTMDGLV 600  
 Qy 601 DLTVGAGCHVLALRSQPLVRKAIEMFNPREVARNVPECNDQVVKGEAGEVRCVLEHVQK 660  
 Db 601 DLTVGAGCHVLALRSQPLVRKAIEMFNPREVARNVPECNDQVVKGEAGEVRCVLEHVQK 660  
 Qy 661 STRDLRREGQIQSVVTVYDLALDSGRPHSAVFNETKNSRRQTQVLGLTQTCETLKLQLP 720  
 Db 661 STRDLRREGQIQSVVTVYDLALDSGRPHSAVFNETKNSRRQTQVLGLTQTCETLKLQLP 720  
 Qy 721 NCIEDPVSPVILRLNFSVGTPLSAFNLRLPVLAEDAQRLLPTALFPPEKNCNDNICODD 780  
 Db 721 NCIEDPVSPVILRLNFSVGTPLSAFNLRLPVLAEDAQRLLPTALFPPEKNCNDNICODD 780  
 Qy 781 LSITTFMSLDCLVWGGPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQORS 840  
 Db 781 LSITTFMSLDCLVWGGPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQORS 840  
 Qy 841 QRSNRLACESASSTEVSGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGNKLLKA 900  
 Db 841 QRSNRLACESASSTEVSGALKSTSCSINHPIIPENSEVTNITFDVDSKASLGNKLLKA 900  
 Qy 901 NVTSENNMPTNKTEFQLELPVKAVYVWVTSHGVSSTKYLNFPTASNTSRVWQHOYQVSN 960  
 Db 901 NVTSENNMPTNKTEFQLELPVKAVYVWVTSHGVSSTKYLNFPTASNTSRVWQHOYQVSN 960  
 Qy 961 LGQSLPISLVFLVRLNQTVINDRPQVTFSENLSTCTKTERLPSSHSDFLAELRKAPV 1020  
 Db 961 LGQSLPISLVFLVRLNQTVINDRPQVTFSENLSTCTKTERLPSSHSDFLAELRKAPV 1020  
 Qy 1021 VNCISIAVCORIQCDIPFPGIOEEFNATLKGMLSDWYIKTSHNHLIVSTAEILFNDVSF 1080  
 Db 1021 VNCISIAVCORIQCDIPFPGIOEEFNATLKGMLSDWYIKTSHNHLIVSTAEILFNDVSF 1080

```
QY 1081 TLLPGCGAFVRSQETKVPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFFKROYKD 1140
DB 1081 TLLPGCGAFVRSQETKVPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFFKROYKD 1140
QY 1141 MMSEGGPPGABPQ 1153
DB 1141 MMSEGGPPGABPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 100.0%; Score 5956; DB 3; Length 1153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRVLLLTALTICGPNLDTENAMTFQENARFGQSVQLOGSRVVGAPQEIIVANQR 60
DB 1 MALRVLLLTALTICGPNLDTENAMTFQENARFGQSVQLOGSRVVGAPQEIIVANQR 60
QY 61 GSLYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATSPQLLACGPTVHQTCSENTYVK 120
DB 61 GSLYQCDYSTGSCPIRLQVPEAVNMSLGLSLAATSPQLLACGPTVHQTCSENTYVK 120
QY 121 GLCFLGSLNRQOPQKPEALRCQPEDSDIAFLIDSGSIIIPDPRMKPEFVSTWEOQL 180
DB 121 GLCFLGSLNRQOPQKPEALRCQPEDSDIAFLIDSGSIIIPDPRMKPEFVSTWEOQL 180
QY 181 KSKTLFSLMAYSEEPRIHPTFKFQNNPNRSLVXPIITQLLGRTHATGIRKVVRELFN 240
DB 181 KSKTLFSLMAYSEEPRIHPTFKFQNNPNRSLVXPIITQLLGRTHATGIRKVVRELFN 240
QY 241 ITWGARKNAKILWITDGEKFGDPLGYEIVIPADREGVIRVIGVDAFRSEKROEL 300
DB 241 ITWGARKNAKILWITDGEKFGDPLGYEIVIPADREGVIRVIGVDAFRSEKROEL 300
QY 301 NTIASKPRDHVQVNNFEALKTIQNLREKIPIAECTQTGSSSSPEHEMSQSGPSAAIT 360
DB 301 NTIASKPRDHVQVNNFEALKTIQNLREKIPIAECTQTGSSSSPEHEMSQSGPSAAIT 360
QY 361 SNGELLSTVGSYDWAGGVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAILLNRVQSLV 420
DB 361 SNGELLSTVGSYDWAGGVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAILLNRVQSLV 420
QY 421 LGAPRYQHI GLVAFRONTGWMESNANVKGTQICAYFGASLCSDVDVDSNGSTDVLIGAP 480
DB 421 LGAPRYQHI GLVAFRONTGWMESNANVKGTQICAYFGASLCSDVDVDSNGSTDVLIGAP 480
QY 481 HYYEQTRGGQVSVCPPLPRQARWQCDVAVLYGCGQFWGRFGAALTIVLGDVNGDKLTDVA 540
```

```
DB 481 HYYEQTRGGQVSVCPPLPRQARWQCDVAVLYGCGQFWGRFGAALTIVLGDVNGDKLTDVA 540
QY 541 IGAPGEEDNRCAVYLFHGTSGSGISPSHSQRIAGSKLSRLOFTQGSLSGGQDLTMDGLV 600
DB 541 IGAPGEEDNRCAVYLFHGTSGSGISPSHSQRIAGSKLSRLOFTQGSLSGGQDLTMDGLV 600
QY 601 DLTVGAQGHVLLRSQPVLRUKAIMEFNPREVARNVFECDNVVKGKAGVRVCLHVOK 660
DB 601 DLTVGAQGHVLLRSQPVLRUKAIMEFNPREVARNVFECDNVVKGKAGVRVCLHVOK 660
QY 661 STRDLREGQIQSVVTVYDLALDSCRPHSRVAFNFTKNSRQOTQVLGLTQTCETLKLQLP 720
DB 661 STRDLREGQIQSVVTVYDLALDSCRPHSRVAFNFTKNSRQOTQVLGLTQTCETLKLQLP 720
QY 721 NCIEDPVSPIVLRNPLSLVGTPLSAPGNLRDVLAEADQRLFTALFPPEKNCNDNICODD 780
DB 721 NCIEDPVSPIVLRNPLSLVGTPLSAPGNLRDVLAEADQRLFTALFPPEKNCNDNICODD 780
QY 781 LSITFSFMSLDCLVVGSPREENVTVTVRNDCGDSVRTQVTFPPFLDLSYRKVSTLQNR 840
DB 781 LSITFSFMSLDCLVVGSPREENVTVTVRNDCGDSVRTQVTFPPFLDLSYRKVSTLQNR 840
QY 841 QRSWLACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKA 900
DB 841 QRSWLACESASSTEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKA 900
QY 901 NVTSENMPRTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
DB 901 NVTSENMPRTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSN 960
QY 961 LGQRLSLPISLVLVPLVRLNQTVIWRDPQVTSSEMLSSSTCHTKERLPSHSDFLAELRKAPV 1020
DB 961 LGQRLSLPISLVLVPLVRLNQTVIWRDPQVTSSEMLSSSTCHTKERLPSHSDFLAELRKAPV 1020
QY 1021 VNCSTAVCORIQCDIPFFGIQEEFNATLKGMLSDWYIKTSHNHLIIVSTAELIFNDSVF 1080
DB 1021 VNCSTAVCORIQCDIPFFGIQEEFNATLKGMLSDWYIKTSHNHLIIVSTAELIFNDSVF 1080
QY 1081 TLLPQCGAFVRSQETKVPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFFKROYKD 1140
DB 1081 TLLPQCGAFVRSQETKVPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFFKROYKD 1140
QY 1141 MMSEGGPPGABPQ 1153
DB 1141 MMSEGGPPGABPQ 1153

RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
```

; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-688-307A-3									
Query Match 100.0%; Score 5956; DB 4; Length 1153;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MALEVLILLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR	60						
Db	1	MALEVLILLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR	60						
Qy	61	GSLYQCDYSTGSCPEIRLQVPEAVNMSGLSLAATTSPPQLLACGPTVHQTCSNTYVK	120						
Db	61	GSLYQCDYSTGSCPEIRLQVPEAVNMSGLSLAATTSPPQLLACGPTVHQTCSNTYVK	120						
Qy	121	GLCFLFGSNLRQOPKQPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEFVSTVMEQL	180						
Db	121	GLCFLFGSNLRQOPKQPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEFVSTVMEQL	180						
Qy	181	KSKTFLSLMOYSEEFRIHFTTFKEFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN	240						
Db	181	KSKTFLSLMOYSEEFRIHFTTFKEFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN	240						
Qy	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVGDAPRSEKSRQEL	300						
Db	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVGDAPRSEKSRQEL	300						
Qy	301	NTIASKPRDHVQVNNFEALKTIONLREKIPAIETGTQTGSSSSPHEMSQSGFSAIT	360						
Db	301	NTIASKPRDHVQVNNFEALKTIONLREKIPAIETGTQTGSSSSPHEMSQSGFSAIT	360						
RESULT 10									
US-09-350-259-3									
; Sequence 3, Application US/09350259									
; Patent No. 6620915									
; GENERAL INFORMATION:									
; APPLICANT: Gallatin, Michael W.									
; APPLICANT: Van der Vieren, Monica									
; TITLE OF INVENTION: No. 6620915el Human 2									
; FILE REFERENCE: 27866/35004									
; CURRENT APPLICATION NUMBER: US/09/350,259									
; EARLIER FILING DATE: 1999-07-08									
; EARLIER APPLICATION NUMBER: 09/193,043									
; EARLIER FILING DATE: 1998-11-16									
; EARLIER APPLICATION NUMBER: 08/173,497									
; EARLIER FILING DATE: 1993-12-23									
; EARLIER APPLICATION NUMBER: 08/286,889									
; EARLIER FILING DATE: 1994-08-05									
; EARLIER APPLICATION NUMBER: 08/362,652									
; EARLIER FILING DATE: 1994-12-21									
; EARLIER APPLICATION NUMBER: 08/943,363									
; EARLIER FILING DATE: 1997-10-03									
; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: Patentin Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-350-259-3									
Query Match 100.0%; Score 5956; DB 4; Length 1153;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MALEVLILLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR	60						
Db	1	MALEVLILLTALTLCHGFNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVANOR	60						
Qy	61	GSLYQCDYSTGSCPEIRLQVPEAVNMSGLSLAATTSPPQLLACGPTVHQTCSNTYVK	120						
Db	61	GSLYQCDYSTGSCPEIRLQVPEAVNMSGLSLAATTSPPQLLACGPTVHQTCSNTYVK	120						
Qy	121	GLCFLFGSNLRQOPKQPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEFVSTVMEQL	180						
Db	121	GLCFLFGSNLRQOPKQPFPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEFVSTVMEQL	180						
Qy	181	KSKTFLSLMOYSEEFRIHFTTFKEFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN	240						
Db	181	KSKTFLSLMOYSEEFRIHFTTFKEFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFN	240						
Qy	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVGDAPRSEKSRQEL	300						
Db	241	ITNGARKNAFKILVITDGEKFGDPLGYEDVPEADREGVIRVYVIGVGDAPRSEKSRQEL	300						
Qy	301	NTIASKPRDHVQVNNFEALKTIONLREKIPAIETGTQTGSSSSPHEMSQSGFSAIT	360						
Db	301	NTIASKPRDHVQVNNFEALKTIONLREKIPAIETGTQTGSSSSPHEMSQSGFSAIT	360						
Qy	361	SNGLLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILNRNVOSLV	420						
Db	361	SNGLLSTVGSYDWAGGVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILNRNVOSLV	420						
Qy	421	IGAPRYOHIGLVAFRONTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAP	480						
Db	421	IGAPRYOHIGLVAFRONTGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAP	480						
Qy	481	HYETQTRGGQSVVCLPRGQARWQCDAVLYGEQGPWGRFGAALTVLGVDNGDKLTDVA	540						
Db	481	HYETQTRGGQSVVCLPRGQARWQCDAVLYGEQGPWGRFGAALTVLGVDNGDKLTDVA	540						
Qy	541	IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLV	600						
Db	541	IGAPGEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLV	600						
Qy	601	DLTVGAQGHVLLLSQPVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQK	660						
Db	601	DLTVGAQGHVLLLSQPVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQK	660						
Qy	661	STRDLRREGQTSQVVTYDLALDSGRPHSRVFNFTKSTRQTOVLGLTQTCETLKLQLP	720						
Db	661	STRDLRREGQTSQVVTYDLALDSGRPHSRVFNFTKSTRQTOVLGLTQTCETLKLQLP	720						
Qy	721	NCIEDPVSPIVLRNFSLVGTPLSARGLRPLVLAEDAQRUFTALFPFKKNGNDNICDD	780						
Db	721	NCIEDPVSPIVLRNFSLVGTPLSARGLRPLVLAEDAQRUFTALFPFKKNGNDNICDD	780						
Qy	781	LSITFSFMSLCLVVGPPREFNVTVVRNDEDSYRTQVTFPPPLDLSYRKVSTLQNR	840						
Db	781	LSITFSFMSLCLVVGPPREFNVTVVRNDEDSYRTQVTFPPPLDLSYRKVSTLQNR	840						
Qy	841	QRNRLACASSTEVSGALAKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLJKA	900						
Db	841	QRNRLACASSTEVSGALAKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLJKA	900						
Qy	901	NVTSENMPRTNKTEFQLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHOYQVSN	960						
Db	901	NVTSENMPRTNKTEFQLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHOYQVSN	960						

Db 301 NTIASKPRDHVQVNNFEALKTIONQLREKIFAIEGTQTGSSSSPEHEMSQSGFSAIT 360  
Qy 361 SNGPILLSVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLV 420  
Db 361 SNGPILLSVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLV 420  
Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAP 480  
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAP 480  
Qy 481 HYEQTRGGQSVSCPLPRGQARMQCDVLYGEOQPMGRFGAALTVLGVDNGDKLTDVA 540  
Db 481 HYEQTRGGQSVSCPLPRGQARMQCDVLYGEOQPMGRFGAALTVLGVDNGDKLTDVA 540  
Qy 541 IGAPEDENRGAIVLPHGTSGSGTSPSHSQRIAGSKLSPRLOYGOSLGGQDITMDGLV 600  
Db 541 IGAPEDENRGAIVLPHGTSGSGTSPSHSQRIAGSKLSPRLOYGOSLGGQDITMDGLV 600  
Qy 601 DLTGAQGHVLLRSQVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660  
Db 601 DLTGAQGHVLLRSQVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660  
Qy 661 STRDLREGQIQSVVYDIALDSGRPHSRVFNETHKSTRTOVGLGTQTCETLKLQLP 720  
Db 661 STRDLREGQIQSVVYDIALDSGRPHSRVFNETHKSTRTOVGLGTQTCETLKLQLP 720  
Qy 721 NCIDPVSPIVLRNLSVGTPLSAFGNLRVLAEDAQRLEFTALFPPEKNCNDNIQDD 780  
Db 721 NCIDPVSPIVLRNLSVGTPLSAFGNLRVLAEDAQRLEFTALFPPEKNCNDNIQDD 780  
Qy 781 LSITFSMSLCLVVGGRPRENVTVVNDGEDSYRTQVTFPPPLDLSYRKVSTLQORS 840  
Db 781 LSITFSMSLCLVVGGRPRENVTVVNDGEDSYRTQVTFPPPLDLSYRKVSTLQORS 840  
Qy 841 QRSRLACESASTEVSGALKSTCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900  
Db 841 QRSRLACESASTEVSGALKSTCSINHPIFPENSEVTNITFDVDSKASLGNKLLKA 900  
Qy 901 NVTSENNMPRTKTEFOLELPVKAVVMTVSHGVSTKYLNAFTASENTRVMAHQYQVSN 960  
Db 901 NVTSENNMPRTKTEFOLELPVKAVVMTVSHGVSTKYLNAFTASENTRVMAHQYQVSN 960  
Qy 961 LGQSLPLSLVFLVPLVRLNQVIVDRPOVTSENLSTCHTKERLPSHSDPLAELRAPV 1020  
Db 961 LGQSLPLSLVFLVPLVRLNQVIVDRPOVTSENLSTCHTKERLPSHSDPLAELRAPV 1020  
Qy 1021 VNCISAVCQRIQCDIPPFQIGIEEFNATLKGNLSDFDWIKTSHNHLIIVSTAEILFNDVSF 1080  
Db 1021 VNCISAVCQRIQCDIPPFQIGIEEFNATLKGNLSDFDWIKTSHNHLIIVSTAEILFNDVSF 1080  
Qy 1081 TLLPQGNFVBSOTETKVEPSEVNPPLIVGSSVGGILLALITAAALYKLGPEKQYKD 1140  
Db 1081 TLLPQGNFVBSOTETKVEPSEVNPPLIVGSSVGGILLALITAAALYKLGPEKQYKD 1140  
Qy 1141 MWSEGGPGGABPQ 1153  
Db 1141 MWSEGGPGGABPQ 1153

RESULT 11

US-08-476-062A-43  
; Sequence 43, Application US/08476062A  
; Patent No. 5877275  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY  
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; STREET: Fish & Richardson P.C.  
; CITY: Boston  
; STATE: MA

COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,062A  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/216,081  
FILING DATE: 21-MAR-1994  
APPLICATION NUMBER: 07/637,830  
FILING DATE: 04-JAN-1991  
APPLICATION NUMBER: 07/539,842  
FILING DATE: 18-JUN-1990  
APPLICATION NUMBER: 07/212,573  
FILING DATE: 28-JUN-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/068003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1152 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-476-062A-43

Query Match 99.5%; Score 5925.5; DB 2; Length 1152;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
Qy 1 MALRVLALLTALTLCHGNLDENAMTFQENARGFGQSVVQLQGRVVGAPQBIIVANQR 60  
Db 1 MALRVLALLTALTLCHGNLDENAMTFQENARGFGQSVVQLQGRVVGAPQBIIVANQR 60  
Qy 61 GSYQCDYTGSCPTRLQVPEAVNMSGLSLAATTSPLLACQPTVHCTSENYUK 120  
Db 61 GSYQCDYTGSCPTRLQVPEAVNMSGLSLAATTSPLLACQPTVHCTSENYUK 120  
Qy 121 GLCFLPGSNLRQOPKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMRKEFVSTVMEQL 180  
Db 121 GLCFLPGSNLRQOPKPFPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMRKEFVSTVMEQL 180  
Qy 181 KSKTTLFSLMQYSBEPRIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240  
Db 181 KSKTTLFSLMQYSBEPRIHPTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFN 240  
Qy 241 ITNGARKNAPKILVWITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300  
Db 241 ITNGARKNAPKILVWITDGEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKSRQEL 300  
Qy 301 NTIASKPRDHVQVNNFEALKTIONQLREKIFAIEGTQTGSSSSPEHEMSQSGFSAIT 360  
Db 301 NTIASKPRDHVQVNNFEALKTIONQLREKIFAIEGTQTGSSSSPEHEMSQSGFSAIT 360  
Qy 361 SNGPILLSVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLV 420  
Db 361 SNGPILLSVGSYDMAGGVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLV 420  
Qy 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAP 480  
Db 421 LGAPRYQHIGLVAMFRQNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAP 480  
Qy 481 HYEQTRGGQSVSCPLPRGQARMQCDVLYGEOQPMGRFGAALTVLGVDNGDKLTDVA 540

481 HYTEQTRGGQSVCLPRG-RARWQCDVLYGQGPWGRFGAALTVLGDVNGKLTJVA 539  
541 IGAPGEDNRGAVYLFHGTSGSISHSORIRAGSKLSPRLQYFGQSLSGGQSLTMDGLV 600  
540 IGAPGEDNRGAVYLFHGTSGSISHSORIRAGSKLSPRLQYFGQSLSGGQSLTMDGLV 599  
601 DLTVGAQGHVLLRSQPLVRKXAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660  
600 DLTVGAQGHVLLRSQPLVRKXAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 659  
661 STRDLRREGQIQSVVTVYDIALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQLP 720  
660 STRDLRREGQIQSVVTVYDIALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQLP 719  
721 NCTEDPVSPIVLRNLFNSLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICODD 780  
720 NCTEDPVSPIVLRNLFNSLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICODD 779  
781 LSITPFSNLSLDCLVGSGPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNR 840  
780 LSITPFSNLSLDCLVGSGPREFNTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNR 839  
841 QRSWLACSSASTSVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 900  
840 QRSWLACSSASTSVSGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKA 899  
901 NVTSENMPRTNKTEQLELPVKIAYVYVTVTSHGVSTKYNFTASENTSRVQHQVQVSN 960  
900 NVTSENMPRTNKTEQLELPVKIAYVYVTVTSHGVSTKYNFTASENTSRVQHQVQVSN 959  
961 LQORSIPISLFLVPLVRLNQTWIDRPQVTFPSENLSSTCHTERLPSSHDFLAEARKAPV 1020  
960 LQORSIPISLFLVPLVRLNQTWIDRPQVTFPSENLSSTCHTERLPSSHDFLAEARKAPV 1019  
1021 VNCSTAVCQRIQCDIPFFGQIEEFNATLXGNLSFDWYIKTSHNHLIIVSTABILFNDVSF 1080  
1020 VNCSTAVCQRIQCDIPFFGQIEEFNATLXGNLSFDWYIKTSHNHLIIVSTABILFNDVSF 1079  
1081 TLLPGGAFVRSQTEKTEKPEFVNPPLPLIVGSSVGGILLALLIITAALYKLGFPPKQYKD 1140  
1080 TLLPGGAFVRSQTEKTEKPEFVNPPLPLIVGSSVGGILLALLIITAALYKLGFPPKQYKD 1139  
1141 MMESEGGPPCAEPQ 1153  
1140 MMESEGGPPCAEPQ 1152

RESULT 12  
PCT-US96-01314-43  
Sequence 43, Application PC/TUS9601314  
GENERAL INFORMATION:  
APPLICANT: M. Amin Arnaout  
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN  
TITLE OF INVENTION: ANTAGONISTS  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01314  
FILING DATE: 30-JAN-96  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/380,167

FILING DATE: 30-JAN-95  
ATTORNEY/AGENT INFORMATION:  
NAME: John W. Freeman  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/267001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1152  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
PCT-US96-01314-43  
Query Match 99.5%; Score 5925.5; DB 5; Length 1152;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1150; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 1 MALRVLLLTALTTLCHGFNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPOEIVAAQR 60  
DB 1 MALRVLLLTALTTLCHGFNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAPOEIVAAQR 60  
QY 61 GSLYQCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120  
DB 61 GSLYQCDYSTGSCBPIRLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVK 120  
QY 121 GLCFLGSLNLRQQQKPPREALRGCPQSDSLAFILDKSGSIIBHDFRMRKPFVSTVMSQL 180  
DB 121 GLCFLGSLNLRQQQKPPREALRGCPQSDSLAFILDKSGSIIBHDFRMRKPFVSTVMSQL 180  
QY 181 KSKTLPFLSMQYSSEFRIHFTFKFQNNPNRSLVKPITQLLGRTHHTATGIRKVVRELFN 240  
DB 181 KSKTLPFLSMQYSSEFRIHFTFKFQNNPNRSLVKPITQLLGRTHHTATGIRKVVRELFN 240  
QY 241 ITTGARKNAFKILVITDGEKFGDPLGYEDVTPADRGVTRVYVIGCDARSEKSRQEL 300  
DB 241 ITTGARKNAFKILVITDGEKFGDPLGYEDVTPADRGVTRVYVIGCDARSEKSRQEL 300  
QY 301 NTIASKPPDRHVFOVNNPFEALKTIONQLREKIIPAEIGTQTCSSSSFEHMSQEGFSAIT 360  
DB 301 NTIASKPPDRHVFOVNNPFEALKTIONQLREKIIPAEIGTQTCSSSSFEHMSQEGFSAIT 360  
QY 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAAILRNRVQSLV 420  
DB 361 SNGPLLSTVGSYDWAGGVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAAILRNRVQSLV 420  
QY 421 LGAPRYQHILGVAMFRONTGWESNANVKGTOIGAYFGASLCSVDMSNGSTDLVLIGAP 480  
DB 421 LGAPRYQHILGVAMFRONTGWESNANVKGTOIGAYFGASLCSVDMSNGSTDLVLIGAP 480  
QY 481 HYTEQTRGGQSVCLPRG-RARWQCDVLYGQGPWGRFGAALTVLGDVNGKLTJVA 540  
DB 481 HYTEQTRGGQSVCLPRG-RARWQCDVLYGQGPWGRFGAALTVLGDVNGKLTJVA 539  
QY 541 IGAPCEDNRGAVYLFHGTSGSISHSORIRAGSKLSPRLQYFGQSLSGGQSLTMDGLV 600  
DB 541 IGAPCEDNRGAVYLFHGTSGSISHSORIRAGSKLSPRLQYFGQSLSGGQSLTMDGLV 599  
QY 601 DLTVGAQGHVLLRSQPLVRKXAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 660  
DB 601 DLTVGAQGHVLLRSQPLVRKXAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQK 659  
QY 661 STRDLRREGQIQSVVTVYDIALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQLP 720  
DB 661 STRDLRREGQIQSVVTVYDIALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQLP 719  
QY 721 NCTEDPVSPIVLRNLFNSLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICODD 780  
DB 721 NCTEDPVSPIVLRNLFNSLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICODD 779



```

/ COUNTRY: US
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ OPERATING SYSTEM: Windows95
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/476,662A
/ FILING DATE: 07-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/216,081
/ FILING DATE: 21-MAR-1994
/ APPLICATION NUMBER: 07/637,830
/ FILING DATE: 04-JAN-1991
/ APPLICATION NUMBER: 07/539,842
/ FILING DATE: 18-JUN-1990
/ APPLICATION NUMBER: 07/212,573
/ FILING DATE: 28-JUN-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Freeman, John W.
/ REGISTRATION NUMBER: 29,066
/ REFERENCE/DOCKET NUMBER: 00786/068003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 44:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1163 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-476-062A-44

Query Match 58.8%; Score 3503; DB 2; Length 1163;
Best Local Similarity 61.3%; Pred. No. 3.8e-287;
Matches 700; Conservative 138; Mismatches 297; Indels 6; Gaps 4;

QY 5 VLLTALTLCGNNLTENAMFOENARGGVQSVQLOQSRVVGAPQBIIVAAORGSLY 64
DB 8 LLLFTALSLGSLNLTBELTAPRVDSAGFSVQVYANSVWVGAPQKITAANQTGGLY 67
QY 65 QCDYSTGSCPTLRQVAVANNSLGLSAAITSPQLLACGPTVHQTCSNTYVKGICF 124
DB 68 QCCYSTGACGCEPIGLQVPPAVANNSLGLSLASTTSPSLLACGPTVHBCGRNMLTGLCF 127
QY 125 LFGSNLRQCPQKPEALRCGPQSDSDIAFLIDGSGSIIIPHDFRMRKPFVSTVMBOLKSK 184
DB 128 LQPT--QTLRLVSRQCPQEQDQIVFLIDGSGSISSENATPMQFVRAVISQFRPS 185
QY 185 TLFLSMQYSEFRIHTPFKEQNNPNRSLVXPIITQLGRHTATGIRKVVRELFINITNG 244
DB 186 TQFSLMQFSNKKFQHTPFEEFRSTNPLSLASVHQLQGFTVATATQVWVHRLPHASYG 245
QY 245 ARKNAEKLIVITDGKFGDPIGYDVIVPEADREGVIRYVIGVDAPFRSEKSRQSLNTIA 304
DB 246 ARDATKILIVITDGKFGDSDLDYKDVIVPAARAAGIIRYVIGVLAQFQNRNKNELNDIA 305
QY 305 SKPPDRHVQVNNFEALTIQNLREKIPIAEGTQTGSSSFEHMSQEGFSAALITNSGP 364
DB 306 SKPSQHEIKVEDFDALQDKIQNLKEKIPIAEGTETITSSSSFELEMAQSGFSAVTPDGP 365
QY 365 LLSVGSYDWAGGVLYTSEKSTFTNTRVDSMDNDAYLGVAALILNRVQSLVGLAP 424
DB 366 VLGAVGSPFTWGGGAFIYPPNMSFTFINNSQENVMDRDSYLGYSTELALWKVQSLVGLAP 425
QY 425 RYCHIGLVAMFRONTGWESNANVKQTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYE 484
DB 426 RYQHTGKAVIPTQVSQWKKAEVTTQIGSIFGASLCSVDVDTGSDLDVLVIGAPHYYE 485
QY 485 QTRGGQVSVCPPLRGORARQCDAVLVYGEQGPWGFAGALTIVLGDVNGDKLTDVAIGAP 544

```

```

DB 486 QTRGGQVSVCPPLRGWR-RWWDVAVLVGEQHPGRFGAALTIVLGDVNGDKLTDVIGAP 544
QY 545 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFOOSLSGGODLTMGDLVLT 604
DB 545 GEENRGAVYLFHGVLGSPISPSHSQRIAGSKLSPRLQYFOQALSGGODLTQDGLVDLAV 604
QY 605 GAQGHVLLRSQPVLRVKAIMSENPREVARNVPCNDQVWKGKEAGEVVRVCLHVQKSTRD 664
DB 605 GARGQVLLRTRPVLWGVSVNQFIPAIIPRPAFCREQVQVSEQTLVQSNICLIYDKRSKN 664
QY 665 RLREGQIQSVVYDIALDGSGRPHSRVFNPNSTRQQTQVGLGTQTCTETKLQLPNCIE 724
DB 665 LLSRDLQSSVTLDLALDPGRSLSPRATFOETKNSLSRVRVLGKAKHCENFNLLPSCVE 724
QY 725 DPVSPVLRNLNFSLVGTPLSAFGLRPLVLAEDAORLFTALPPEKNCNDNI CODDLSIT 784
DB 725 DSVTPITLRLNFTLVGKPELLAFRLNRLMAMLAQRYFTASLPPEKNCADHI CQNLGIS 784
QY 785 FSNMSLQCLVYGGPRBFNVTVVRNDGEDSYRTQVTFPFPLDLSYRKVYSTLQNRQSRSW 844
DB 785 FSPFLKSLVGSNLELMAEVMVWVNDGEDSYGTTITFSPAGLSYRYVAEGQKQQLASL 844
QY 845 PLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTS 904
DB 845 HLTCDSPVPG--SQGTWSTSCRIHELIFRGAQITFLATFDVSPKAVILGDRLLLTANVSS 902
QY 905 ENNMPTNKTEFQLELPVKYAVYVMTVSHGVSTKYLNPFTAS-ENTSRVMQHQYQVSNLQ 963
DB 903 ENTPTSKITTFQLELPVKYAVYTVSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQ 962
QY 964 RSLPISLVLPVRLNQTVIWDRCQVTFSENLSTCTHKERLPSHSDFLAELRKAPVNC 1023
DB 963 RDLFVSINFVWVVELNQEAVMVDEVSHPPQNPRLCSKSEKLIAPPASDFLAHIQKNPVLDC 1022
QY 1024 SIACQRIQCDIPFFGIEEFENATLKNLSFDMWIKTSHNLLIVSTAEILFNDSVFLL 1083
DB 1023 SIACGLRFRCDVPFSVQSELDFTLKNLSFGNVRQILQKVVSVVVAEITFDTSVYQL 1082
QY 1084 PQGAFVRSQTEKTVPEPEVNPPLPLIVGSSVGGLLLALITALYKUGFKRQYKDMMS 1143
DB 1083 PQGAPEGAQTTVLEKYKHNPPTPLIVGSSIGGLLALITALYKVGFFKRYKEMME 1142
QY 1144 E 1144
DB 1143 E 1143

RESULT 15
PCT-US96-01314-44
; Sequence 44, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: June 7, 2004, 17:05:16 ; Search time 14.559 Seconds  
(without alignments)  
7512.163 Million cell updates/sec

Title: US-09-902-481B-3  
Perfect score: 5879  
Sequence: 1 FNLDTENAMTQENARGPGQ.....FKRYQKDMXSGGPPGAEFQ 1137  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Piri:\*  
2: Pir2:\*  
3: Pir3:\*  
4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	1 RWHU1B	cell surface glyco
2	4447	75.6	1153	2 S00551	leukocyte surface
3	3456	58.3	1163	1 RWHU1C	cell surface glyco
4	1532.5	26.1	1170	2 S03308	cell surface glyco
5	1516.5	25.8	1163	2 I56126	lymphocyte fuction
6	1128	19.2	1179	2 A53213	integrin alpha-E c
7	1085.5	18.5	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45914	integrin alpha-2 s
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1054	17.9	1181	2 A33998	integrin alpha-2 c
11	1049	17.8	1180	2 A35854	integrin alpha-1 c
12	866	11.3	1039	2 A41131	lymphocyte-Peyer's
13	842	10.9	1038	2 S05046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha cha
16	579.5	9.9	1054	2 JC7294	alphanp integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.4	1053	2 S44250	integrin alpha-5 c
20	551.5	9.4	1034	2 A36108	integrin alpha-V c
21	539	9.2	1044	2 T10050	integrin alpha-v c
22	533.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	fibronectin recept
24	531	9.0	1073	2 B36429	integrin alpha-6 c
25	530	9.0	1048	2 A27421	integrin alpha-5 c
26	529.5	9.0	1051	2 A40021	integrin alpha-3 c
27	524.5	8.9	1091	2 A41543	integrin alpha-6 c
28	514.5	8.8	1044	2 S16516	integrin alpha-8 c
29	497	8.5	1394	2 A29637	position-specific

RESULT 1

RWU1B

cell surface glycoprotein CD11b precursor [validated] - human  
N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein 1  
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000  
C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; 152567  
R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.  
J. Biol. Chem. 263, 12403-12411, 1988  
A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, B.  
A:Reference number: A31108; MUID:88315033; PMID:2457584  
A:Accession: A31108  
A:Molecule type: mRNA  
A:Residues: 1-1153 <COR>  
A:Cross-references: GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148  
A:Note: part of this sequence was confirmed by protein sequencing  
R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.  
J. Cell Biol. 106, 2153-2158, 1988  
A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor  
A:Reference number: A28915; MUID:88257215; PMID:2454931  
A:Accession: A28915  
A:Molecule type: mRNA  
A:Residues: 1-499,501-965, '967-1153 <ARN>  
A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA5:  
A:Note: the authors translated the codon TAC for residue 1129 as Thr  
A:Note: part of this sequence, including the amino end of the mature protein, was conf:  
R:Shelley, C.S.; Arnaout, M.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991  
A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally r  
A:Reference number: A41600; MUID:92073318; PMID:1683702  
A:Accession: A41600  
A:Molecule type: DNA  
A:Residues: 1-9 <SHE>  
A:Cross-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215  
R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.  
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988  
A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhe:  
A:Reference number: A94193; MUID:88190151; PMID:2833753  
A:Accession: A30892  
A:Molecule type: mRNA  
A:Residues: 917-1042 <AR2>  
A:Cross-references: GB:M18044  
R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.  
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989  
A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recept<  
A:Reference number: A32218; MUID:89098893; PMID:2563162  
A:Accession: A32218  
A:Molecule type: mRNA  
A:Residues: 9-1153 <HTC>  
A:Cross-references: GB:J04145; NID:G189068; PIDN:AAA59503.1; PID:G386975

ALIGNMENTS

integrin - fruit f  
integrin alpha-2b  
integrin alpha-7 c  
alpha-7 integrin -  
glycoprotein IIB -  
leukocyte adhesion  
integrin alpha cha  
F54F2.1 protein -  
integrin alpha v c  
hypothetical prote  
integrin alpha cha  
glycoprotein IIB -  
integrin alpha cha  
glycoprotein IIB -  
integrin alpha-1 -

A>Note: part of this sequence was confirmed by protein sequencing  
 J. Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.  
 J. Immunol. 150, 480-490, 1993  
 A>Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in  
 n during evolution.  
 A:Reference number: A46526; MUID:93123748; PMID:8419480  
 A:Accession: A46526  
 A>Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-499,501-1153 <FILE>  
 A:CROSS-references: GB:IS52227; NID:G263047; PIDN:AAB24821.1; PID:G263049  
 A>Note: the last three bases of intron 13, CAG, are included in some but not all mature  
 A>Note: sequence extracted from NCBI backbone (NCBIP:121963)  
 R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.  
 Biochim. Biophys. Acta 874, 368-371, 1986  
 A>Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp  
 A:Reference number: A90664; MUID:87076671; PMID:3539202  
 A:Accession: A26091  
 A:Molecule type: protein  
 A:Residues: 17-31 <PIE>  
 A:Experimental source: granulocytes  
 R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.  
 Blood 79, 865-870, 1992  
 A>Title: Characterization of the myeloid-specific CD11b promoter.  
 A:Reference number: 152567; MUID:92144986; PMID:1346576  
 A:Accession: 152567  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-9 <RES>  
 A:CROSS-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:9553219  
 A:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
 C:Genetics:  
 A:Gene: GDB:ITGAM; CR3A  
 A:CROSS-references: GDB:120599; OMIM:120980  
 A:Map position: 16p11.2-16p11.2  
 A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites  
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
 C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mac  
 R:1-16/Domain: signal sequence #status predicted <SIG>  
 R:17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>  
 R:17-1108/Domain: extracellular #status predicted <EXT>  
 R:148-318/Domain: von Willebrand factor type A repeat homology <WMA2>  
 F:465-473/Region: calcium/magnesium binding #status predicted  
 F:530-538/Region: calcium/magnesium binding #status predicted  
 F:593-601/Region: calcium/magnesium binding #status predicted  
 F:1109-1134/Domain: transmembrane #status predicted <TM>  
 R:1135-1153/Domain: intracellular #status predicted <INT>  
 F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding  
 Query Match 99.3%; Score 5839; DB 1; Length 1153;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 60  
 Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 76  
 Qy 61 RLQVPEAVNMSLGLSLAATTSPPQLLAGCTVHQTCSNTYVKGCLFGLFSGNLRLQPOK 120  
 Db 77 RLQVPEAVNMSLGLSLAATTSPPQLLAGCTVHQTCSNTYVKGCLFGLFSGNLRLQPOK 136  
 Qy 121 FPARLRCPQEDSDIAFLIDGSGIIPHPFRMKELVSTIMEQLKSKTLFSLIMQYSEEF 180  
 Db 137 FPARLRCPQEDSDIAFLIDGSGIIPHPFRMKELVSTIMEQLKSKTLFSLIMQYSEEF 196  
 Qy 181 RIHFTKEFONNPNASLKPITQLLGRHTATGLKRVRELFNITNGARKNAFKILLL 240  
 Db 197 RIHFTKEFONNPNASLKPITQLLGRHTATGLKRVRELFNITNGARKNAFKILVVI 256  
 Qy 241 TDGEKFDPLGYEDVPEADREGVIRYVIGVGDAFSEKSRQELNTIASKPPRDHVFQVN 300  
 Db 257 TDGEKFDPLGYEDVPEADREGVIRYVIGVGDAFSEKSRQELNTIASKPPRDHVFQVN 316

## RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N:Alternate names: complement-3 receptor alpha chain

C:Species: Mus musculus (house mouse)

C&gt;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 22-Oct-1999

C:Accession: S00551; I59078

R:Pytela, R.

EMBL J. 7, 1371-1378, 1988

A&gt;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the

A:Reference number: S00551; MUID:88312584; PMID:3044779

A:Accession: S00551

A:Molecule type: DNA

A:Residues: 1-1153 &lt;PVT&gt;

A:CROSS-references: EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G52983

A&gt;Note: the authors translated the codon CAC for residue 569 as Gln

R:Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986  
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep  
A:Reference number: 159078; MUID:86287312; PMID:2942940  
A:Accession: 159078  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 11-44 <RES>  
A:Cross-references: GB:M14293; NID:gi98993; PIDN:AAA39484.1; PID:g554193  
C:Genetics:  
C:Gene: Mac-1  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C:Keywords: cell adhesion; glycoprotein; transmembrane protein  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental  
F:148-318/Domain: von Willebrand factor type A repeat homology <VWA>  
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Watch 75.8%; Score 4447; DB 2; Length 1153;  
Best Local Similarity 73.5%; Pred. No. 3.8e-303;  
Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARFGQSVVQLCGSRVVGAPQEIIVANQSGSLYOCXYSTGSCRFI 60  
DB 17 FNLDTEHPTFQENAKGFGQNVVQLGTSVVAAPQEAQVAVQTHGYOCXYSTGRCHPI 76  
QY 61 RLQVPEAVNMSLGLSLAANTSPQLLACGPTVHQCSTENTYVKGCLFLFGSLNLRQPOK 120  
DB 77 PLQVPEAVNMSLGLSLAVSTVPQQLLACGPTVHQCSTENTYVKGCLFLFGSLNLRPQO 136  
QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHPFRWKLAVSTIMEQLKSKTLFSLMQYSEEP 180  
DB 137 FPEALRECHQSSDIIVFLDGGSGINIIDFQMKKEFVSTVMEQFKSKTLFSLMQYSEDF 196  
QY 181 RHFTFKPQNNPNRSLKIPITQLAGRTHTATGLKVVRELFNITNGARKNAKILFLL 240  
DB 197 RHFTFNDKPNPSPSHVSPKQLNGRTKTASGIRKVVRELFHKTNGARENAKILVVI 256  
QY 241 TDEKRGDPLGVEDVITPELDREGVIRYVILGFGDAPRSEKSRQELNVTASKPRDHVFOAN 300  
DB 257 TDEKRGDPLGVEDVITPELDREGVIRYVILGFGDAPRSEKSRQELNVTASKPRDHVFOAN 316  
QY 301 NFEALTKVQNRKELFALEGTOGSSSEFHEMSQEGFSAITNGPLLTSTVGSYDWAG 360  
DB 317 NFEALNTIQNQLKELFALEGTOGSSSEFHEMSQEGFSAITNGPLLTSTVGSYDWAG 376  
QY 361 GVPLTSEKSTFINNTRVDSMDNDAYLGAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420  
DB 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436  
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYECTGCGVSVCPL 480  
DB 437 ENFGTWEPSTKSGSIQIFGASLCSVDVDSNGSTDLVLIGAPHYECTGCGVSVCPL 496  
QY 481 PRGQARWQCDVLYGEOQOPMGRFGAALTVLGDVNGDKLTDVAICAPGEEDNRGAVYLF 540  
DB 497 PRG-RARWQCEALLHGDQHPMGRFGAALTVLGDVNGDKLTDVAICAPGEEDNRGAVYLF 555  
QY 541 HGTSSGSIQSPHSQRTAGSKLSPRLOYFQOSLSGGDLTMDGLVDLTVGAQGHVLLRSQ 600  
DB 556 YGASTASLSASHRIIGAFHFGFLQYFQOSLSGGDLTMDGLVDLTVGAQGHVLLRSQ 615  
QY 601 PVLRKALMEFNPVARNVFCNDQVKGKAGEVRVCLHVQKSTRKLRREGQIQSVVT 660  
DB 616 PVLRLEATWTFSPKVARSVFACQGVLENKQDAGEVRVCLRVKTKLREGDLQSVT 675  
QY 661 YDLALDSGRPHSAVNETNSTRGQTVLGLTQTCETLKLQPLNCIEDPVPVILRLNF 720  
DB 676 YDLALDPVRSIRAFDEFKNNTRRTQVFGLMQKCEKTLKLLPDCVDVDSVPIILRLN 735  
QY 721 SLVGTPLSAFNGNLRPVLAEDAQLFTALFPFPEKNGCNNDICODDSITFSPFMSLCLVVG 780  
DB 736 TLVGEPLASFGNLRPVLANDAQRFPTAMPFFPEKNGCNDSICQDDUSITWSAGLDTLVVG 795

RESULT 3  
RWNUIC  
cell surface glycoprotein CD11c precursor - human  
N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text\_change 22-Jun-1999  
C:Accession: A36584; A35543; S00864  
J:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
J. Biol. Chem. 265, 12750-12751, 1990  
A:Reference number: A36584  
A:Contexts: erratum  
A:Accession: A36584  
A:Molecule type: DNA  
A:Residues: 1-1163 <COR>  
A:Note: this revision to the sequence from reference A35543 includes the carboxyl end  
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
J. Biol. Chem. 265, 2782-2788, 1990  
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.  
A:Reference number: A35543; MUID:90153906; PMID:2303426  
A:Accession: A35543  
A:Molecule type: DNA  
A:Residues: 1-834 <CO2>  
A:Note: this sequence has been revised in reference A36584  
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.  
EMBO J. 6, 4023-4028, 1987  
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte  
A:Reference number: S00864; MUID:88166645; PMID:3327687  
A:Accession: S00864  
A:Molecule type: mRNA  
A:Residues: 1-755, 'L', 757-1163 <CO3>  
A:Cross-references: GB:M61695; EMBL:Y00093; NID:9487829; PIDN:AAA59180.1; PID:g487830  
A:Note: part of this sequence was confirmed by protein sequencing  
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my  
C:Genetics:  
A:Gene: GDB:ITGAX; CD11C  
A:Cross-references: GDB:119758; OMIM:151510  
A:Map position: 16p11.2-16p11.2  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>  
F:210-1107/Domain: extracellular #status predicted <EXT>  
F:1149-319/Domain: von Willebrand factor type A repeat homology <VWA>  
F:1108-1133/Domain: transmembrane #status predicted <TM>  
F:1134-1163/Domain: intracellular #status predicted <INT>  
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

[illegible]









```

QY 956 FLV-----PVELNOTVIWDRPOVTF-SRNL5 980
D 985 LVTHIHOYTKEKNPLLYLTGQDQAGDISCTAEINPLKLPHTA-----PSVSFKNER 1040
QY 981 STCHTERLPSHSDFLAELRKAPVNVCSIAVQRIQCDIPFGIQEFNATLKNLSFDW 1040
D 1041 ---HTKE-----LDCRTTSCSNITCWLKDLKAEYFYNVTVTRVWRT 1080
QY 1041 YIKTSHHLLIVSTAELNDVSFTLLPGOGAFVRSQTETKVPFVDPNPLPIVGVSVG 1100
D 1081 PAASTQTQVLTAAAEIDTNPQLFVIEENAVTIPLMIMKPTKAEVPT--GVIIISIIA 1138
QY 1101 GLLLALITAAALYKLGFFKQYKDM 1125
D 1139 GILLALMTAGLAKLGFKKQYKDM 1163

RESULT 10
A33998
N;Integrin alpha-2 chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 15-Sep-2003
A;Accession: A33998; B56793; A53117
R;Takada, Y.; Henler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: GB:X17033; NID:G33906; PIDN:CAA34894.1; PID:G33907
A;Note: The authors translated the codon GAT for residue 802 as Gin, GTC for residue 803
R;Cattamel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and GPIIc)
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:L24121; NID:G9400342; PIDN:AAA16619.2; PID:G4583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,134,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 17.9%; Score 1054; DB 2; Length 1181;
Best Local Similarity 26.8%; Pred. No. 3 6e-65;
Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;

QY 1 FNLDENATFTQ-ENARGGQSVQL-----QGRVVVGAPQEIIVANORGLYQC--DYST 54
D 30 YNVGPEAKIFSGPSSEQGYAVQVQFINKPKGNLLVGVSPWSPFPPNRMGDVTKCPVDLST 89
QY 55 GSCPEIRLQ-----VPVEAVMISGLSLAATTSPQQLACGPTVHQCSENTVYKGLC 107
D 90 ATCEKLNLTGTSIPNVTEKTNMISGLILTRMGTGGFLTCGPLWAQCGNQYITTVGC 149

```

```

QY 108 FLFGNLRQOPKPFEPALRGCPQEDSDIAFLIDGSGSIIPHFRMKELVSTIMEOLK-- 165
D 150 SDISPDF-QLSASFSPATQCPSSL-IDVVVVCDENSIYPMID--AVKNLEKFEVQGLDIG 205
QY 166 KSKTLPSLMQYSEBERIHFTFKSPONNPNRSLIKPITCLLG-RHTHTATGLAKVRELEN 224
D 206 PKTQVGLLQYANRRVVENLNTYTKEMIVATSTQSYGGDLNTFPATQIYARKYAS 265
QY 225 IYNGARKNAFKILFLTLLTDEKFGDPLGYEDVTEPEIDREGVIRY----VLGF--GDAPRSK 279
D 266 AASGGRSATKVMVVVDTGESH-DGSMKXAVIDQCNDHNLFGIAGVAVLYLNALDITKN 324
QY 280 SPOELNTVASKEPRDHFVQANNFEALKTQVONLRKIPALEGTCTGSSSSPHEMSOEGF 339
D 325 LIKEIKATASIPTEYFPFVNSDEALKEKAGTLGQIFSIETGVGQ-GDNFQEMSVQVGF 383
QY 340 SAAITSNGB--LLSTVSGYDNAGGVPLYSKESKSTFINMT--RVSDMN--DAYLGYAAAI 394
D 384 SADYSSQNDILMLGAVGAFGSGTIVQKTSGLHLIFPKQAFQILQDRNHSYLYGSVAA 443
QY 395 ILNRVQSLVIGAPRYOHILGLVAMPFRONTGWESNANV-----KGTQIGAYFGASLCSV 448
D 444 ISTGESTHPVAGAPRNTYGTQIVLYSVN-----ENGNITVIAHHRGDQIGSYFGVLCSV 498
QY 449 EYDSNGSTDLVLIGAPHYYEQTR--GGQVSVCPLEPRGORARWQCDVAVLYGEQGPWGRFG 506
D 499 DVDKDTITDVLVAGAPMTMSDLKKEEGRYLFTIKKILGQHQ---PLEGPRGIENTRFG 555
QY 507 RALTIVGDNVGDGLTDVAIGARDEEDNEGAVYLFEGTSGSGISPSHSQRIAGS--KLSPR 564
D 556 SAIAALSINDDGFDVTVGSPLENQNSGAVIYVNGHQGT-IRTKYSQIKLSDGDAFRSH 614
QY 565 LQYFQSLSGGQDLTMDGLVLTGVAQGHVLLLSQPLVRKAIEMFPNPREVARVNFBCN 624
D 615 LQYFGRSLDGYDGLNGDSITDVSIGAFQVQVQLWSQSIADVAIEASFTEKI--TLVKN 672
QY 625 DVVVKKEAGEVTVCLHVQKSTRDLREGQIQSVVTVDLALD-----SGRPHSAFVNETK 680
D 673 AQII-----LKLCF-----SAKFRPTQNNQVAIVYNTILDADGFSRVRTSRGLFKENN 721
QY 681 NSTRRQTVLGLTQTC--ETLKLQLPNCIEDVPSPVLRNLFNPSIVGTPLSAPGNLRPVLIA 738
D 722 ERLQKNVNVQASCPHEIIVIEPS---DVVNSLDLREVISLENPGTS-----PALE 772
QY 739 EDAQLFTALPFPKNGCNDNICQDDLSITF-----SFMSLDCLVVGGRPFENVTVTRND 794
D 773 AVSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLXNK 832
QY 795 GEDSVRTQVTFPPLDLSYRKVSTLQORSQORSNRLACESASST-EVSGALKSTSCSINH 853
D 833 RESAINTGIVVDPSNLFF-----ASTSLPVDGTETVTCQVAASQKSAQVADVGY 880
QY 854 PIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEFOLELPVKYAVTMV 913
D 881 PALKREQVTVTFINPDLNLQ-QVQASLSPOALSESQENKADNLVNLKIPLLYDAEI- 938
QY 914 VTHGVSTKYLNPTASENTSNVMQHVQVSNLGR-----SLPISLVFLV----- 958
D 939 ---HLTRSTNINPVEISDGNVPSIVSHSFEDVGPKPIFSLKVTGSEVSMATVHIHPQ 995
QY 959 -----PVLNQTVIWDPRQVTF--SENLSSTCHTKER 988
D 996 YTEKNPLMYLTGVQTDKAGDISCNADINPLKIQGT---SSVSVPKSENF--HTKE- 1047
QY 989 LPSHSDFLAELRKAPVNVCSIAVQRIQCDIPFGIQEFNATLKNLSFDWYIKTSINH 1048
D 1048 -----LNCRTASCNVTKLQVHMKGEYFVNVTVTRINWTPASSTFQT 1091
QY 1049 LLIVSTAEL-LPNDVSFTLLPGOGAFVRSQETKVPFVDPNPLPIVGVSVG 1100
D 1092 VQLTAAAEINTYNPISYVI-----EDNTVTIPLMIMKPEKAEVPGVIGLSIIA 1141

```







Db 963 TVVFEALHNLPRGYVVGWIIAISLLVIGLIFULLAVLWKKMGFFRRRYKEII 1015

RESULT 15

T31437

integrin alpha chain SU2 - sea urchin (lytechinus variegatus)

C:Species: Lytechinus variegatus (variegated urchin)

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000

C:Accession: T31437

R:Hertzler, P.L.; McClay, D.R.

submitted to the EMBL data Library, May 1998

A:Description: Integrin alpha SU2, a sea urchin integrin which binds laminin.

A:Reference number: Z21035

A:Accession: T31437

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1041 <HR>

A:Cross-references: EMBL:AF067658; MID:g3220240; PID:g3220241; PIDN:AAC23572.1

A:Experimental source: developmental stage embryo

C:Function:

A:Description: binds laminin

C:Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;

Best Local Similarity 24.8%; Pred. No. 1.9e-34;

Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAATISNGP--LLSTVSGDWAGVGLYTSKEKSFNNTRVDS-----DMNDAYLG 389

Db 181 QAGFSGILFSDNSALVMPAGSGSYLQGIYVQSLNRSV-VQATQESNTGYSPDNSYRG 239

QY 390 YAAAI--ILNRVQSLVLCAPRYQHI-GLVAMFRONTGMESNANVKGTOIGAYFGASLC 446

Db 240 YSLALGDFNGDGVDQYVGTTPRAESLMGLVAIFDQNLQFN---QVMGTQIVAYFYGSVT 296

QY 447 SYVDVNSGTDLVLIAGPHYEYQTRGGQVSVCPFRGQARQCDAVLVYGEQ----- 498

Db 297 VVDI-NDYDILLVGAPEYMDGPAIQ-----RWEAGAVYYLQNPDPVGPGA 343

QY 499 -----GQWGRFGAALTVLGVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGS 546

Db 344 SNRLSLSSTLIGGQIRSRFGLSIASIGDSNQGFNDVAIGAPYEGDDAGAVYVHG-SAN 402

QY 547 GISPSSHQRIAGKLS-PLQYFGQSLGGDLTMDGLVDLTVGAQ--GHVILLRSQVPL 603

Db 403 GLKSTPAQLTPTLGHSGITTFGSLQGGQMDKKNYFDLLVGAESANTAVLINTRPVV 462

QY 604 RYKAIMENPREVARNFECNDQVYVKGKAGEVRVCLHVOKSTRDLREGQIQSV----- 658

Db 463 SLDATLNTPE-----IGINLENKTYE-LADGTMWTSFIAMT 497

QY 659 -----VTYDLALDSG-RPHSRVFNTRKSTRQTOVLGL-TOTCETLKLQ 702

Db 498 CFTYTGNYLFDHIDISYTVYVDSGIIANRRAMFVNDMSEITKTRRLAVSTQCPDPLRAY 557

QY 703 LPNCIEDPVSPVLRNLNLSVGTPLSAFCN-----LRPVLAEADAQLFTALFPPEK 753

Db 558 VGNSTEDKLTPIKVTLOYDL-----NNDESRLQPHLEIPIIDMATMSTQTKQVSIQN 609

QY 754 NCGNDNICODLSITFPMSGLDCLVGGPREPNTVTVRNDGEDSYRTQVTFPPPLDLSY 813

Db 610 NCVN-NICIPDLADVTV-PNLPNIVIGOTQELTLDVSLNRRGEDAFQSSLSVYPLGIQF 667

QY 814 RYKSTLQHQRSQRWRLACESASSTEVSCALKSTCSINHPIFFEN-----SEVTENIT 867

Db 668 VPL-----ERKANMDSVTCSEDS-----LRITCTGNPMVGKNIIEFGILTSTPQVS 717

QY 868 FVDVS-----KASLGNKLLKANVTSENNMPTNKTEFQLELPVKYAVYVMTVSHGVSTK 922

Db 718 GPKDSIEFFYFKAESSENS--EDPNTLENNELNMTVPVVDCTLKLSASYPEIMVYKSTQED 775

QY 923 YL-----NFTASENTSRVMCHOYQV-----SNLQESLPTS-----LVFLVPV 960

Db 776 YVPPFPKKNASEADIGNEVM-HLYEVRNTGSSNAGEVSLNIQWPKNEDGELYLYLLGI 834

QY 961 RLNOTVIMDRPQ-----VTFSENLSSTCHTKERLPFSHSDFTLAELRKAPVV 1005

Db 835 MTEEGVTCQLTQGGKANPEGVKLEPSTKAKLSNSTTOVSGRKRREFEVAALAQTDN--VI 892

QY 1006 NCSTAVCQRIQCDIPFFGCIQEEFNAT-----LKGNSLFDWYIKTSHNHLIVSTABILF 1059

Db 893 YCASDSCVLINCTI-----DEINASKSKVVRILGRP---W-----ERTF 928

QY 1060 NDSVFTLLPGGAGFVRSTQETKVB--PFEVNP-----LP----- 1092

Db 929 QKAVSELTPVQVQATIASASAAXTIPYNIPLPRDFSDSTRASTLIVTTEELVPPVPTIAW 988

QY 1093 -LIVGSSVGGELLALALTAALYKLGPFKR-----QYKDMMS--EGGPP 1132

Db 989 WIIIVSVLGGIILLIILGLWKGPFERKPKGEEKYAPVASADKDGPP 1038

Search completed: June 7, 2004, 17:17:38

Job time : 19.559 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 9.97194 Seconds  
(without alignments)  
5937.039 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5823.5	99.1	1152	1	ITAM HUMAN
2	4447	75.6	1153	1	ITAM MOUSE
3	3446	58.6	1163	1	ITAD HUMAN
4	3388	57.6	1162	1	ITAD HUMAN
5	1536.5	26.1	1170	1	ITAD HUMAN
6	1516.5	25.8	1163	1	ITAD MOUSE
7	1136.5	19.3	1167	1	ITAD MOUSE
8	1128	19.2	1179	1	ITAD HUMAN
9	1085.5	18.5	1151	1	ITAD HUMAN
10	1079.5	18.4	1189	1	ITAD HUMAN
11	1071	18.2	1170	1	ITAD BOVIN
12	1057	18.0	1178	1	ITAD MOUSE
13	1054	17.9	1181	1	ITAD HUMAN
14	1049	17.8	1180	1	ITAD RAT
15	1042.5	17.7	1167	1	ITAG HUMAN
16	666	11.3	1039	1	ITAD MOUSE
17	642	10.9	1038	1	ITAD HUMAN
18	630	10.7	1035	1	ITAD HUMAN
19	594.5	10.1	1032	1	ITAD XENLA
20	571.5	9.7	1066	1	ITAD CRISP
21	567.5	9.7	1053	1	ITAD MOUSE
22	555.5	9.4	1053	1	ITAD MOUSE
23	551.5	9.4	1034	1	ITAD CHICK
24	542.5	9.2	1050	1	ITAD XENLA
25	539	9.2	1044	1	ITAD MOUSE
26	534.5	9.1	1130	1	ITAD HUMAN
27	533.5	9.1	1072	1	ITAD CHICK
28	532	9.0	1049	1	ITAD HUMAN
29	531.5	9.0	1066	1	ITAD HUMAN
30	530	9.0	1048	1	ITAD HUMAN
31	514.5	8.8	1044	1	ITAD CHICK
32	511.5	8.7	1091	1	ITAD MOUSE
33	505	8.6	1179	1	ITAD MOUSE

#### ALIGNMENTS

##### RESULT 1

ID	ITAM HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnsaut M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnsaut M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	497	8.5	1396	1	ITA2_DROME
35	491.5	8.4	1146	1	ITA1_DROME
36	490	8.3	1039	1	ITAB_HUMAN
37	489	8.3	1025	1	ITAB_HUMAN
38	488	8.3	1033	1	ITAB_MOUSE
39	486	8.3	126	1	ITAM_CAVPO
40	478	8.1	1181	1	ITAM_HUMAN
41	470	8.0	1106	1	ITAT_RAT
42	466	7.9	1226	1	ITAT_CAEEL
43	444.5	7.6	1139	1	ITAT_CAEEL
44	419	7.1	1115	1	ITA3_DROME
45	382	6.5	1000	1	ITA5_DROME

P12080	drosophila
Q24247	drosophila
P08514	homo sapien
P53708	homo sapien
Q93000	mus musculus
P11578	cavia porce
Q13683	homo sapien
Q63258	rattus norv
P34446	caenorhabdi
Q03600	caenorhabdi
O44386	drosophila
Q9wim8	drosophila

SEQUENCE OF 1-9 FROM N.A.  
MEDLINE=92073318; PubMed=1683702;  
Shelley C.S., Arnaout M.A.;  
"The promoter of the CD11b gene directs myeloid-specific and  
developmentally regulated expression.";  
Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).  
[7]  
SEQUENCE OF 1-9 FROM N.A.  
TISSUE=Blood;  
MEDLINE=92144986; PubMed=1346576;  
Pahl H.L., Rosmarin A.G., Tenen D.G.;  
"Characterization of the myeloid-specific CD11b promoter.";  
Blood 79:865-870(1992).  
[8]  
SEQUENCE OF 17-31.  
MEDLINE=87076671; PubMed=3539202;  
Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;  
"N-terminal sequence of human leukocyte glycoprotein Mol:  
conservation across species and homology to platelet Iib/IIia.";  
Biochim. Biophys. Acta 874:368-371(1986).  
[9]  
X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.  
MEDLINE=95171458; PubMed=7867070;  
Lee J.O., Rieu P., Arnaout M.A., Liddington R.;  
"Crystal structure of the A domain from the alpha subunit of integrin  
CR3 (CD11b/CD18).";  
Cell 80:631-638(1995).  
[10]  
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.  
MEDLINE=96363671; PubMed=8747460;  
Lee J.O., Banks L.A., Arnaout M.A., Liddington R.;  
"Two conformations of the integrin A-domain (I-domain): a pathway for  
activation?";  
Structure 3:1333-1340(1995).  
[11]  
X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.  
MEDLINE=98362595; PubMed=9687375;  
Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,  
Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,  
Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,  
Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;  
"Cation binding to the integrin CD11b I domain and activation model  
assessment.";  
Structure 6:923-935(1998).  
[12]  
3D-STRUCTURE MODELING OF 17-616.  
MEDLINE=98226734; PubMed=9560195;  
Oxvig C., Springer T.A.;  
"Experimental support for a beta-propeller domain in integrin alpha-  
subunits and a calcium binding site on its lower surface.";  
Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).  
-!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS  
ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES  
AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.  
IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF  
THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D  
PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR  
FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES  
OF FIBRINOGEN GAMMA CHAIN.  
-!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M  
ASSOCIATES WITH BETA-2.  
-!- SUBCELLULAR LOCATION: PREDOMINANTLY EXPRESSED IN MONOCYTES AND  
GRANULOCYTES.  
-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
-!- SIMILARITY: Belongs to the integrin alpha chain family.  
-!- SIMILARITY: Contains 1 VWFA domain.  
-!- SIMILARITY: Contains 7 FG-GAP repeats.  
-!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;  
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11b.htm".  
-----  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; J03925; AAA59544.1; -  
CC EMBL; M18044; AAA59491.1; -  
CC EMBL; J04145; AAA59903.1; -  
CC EMBL; S52227; AAB24821.1; -  
CC EMBL; S52152; AAB24821.1; JOINED.  
CC EMBL; S52153; AAB24821.1; JOINED.  
CC EMBL; S52154; AAB24821.1; JOINED.  
CC EMBL; S52155; AAB24821.1; JOINED.  
CC EMBL; S52157; AAB24821.1; JOINED.  
CC EMBL; S52159; AAB24821.1; JOINED.  
CC EMBL; S52161; AAB24821.1; JOINED.  
CC EMBL; S52164; AAB24821.1; JOINED.  
CC EMBL; S52165; AAB24821.1; JOINED.  
CC EMBL; S52167; AAB24821.1; JOINED.  
CC EMBL; S52169; AAB24821.1; JOINED.  
CC EMBL; S52170; AAB24821.1; JOINED.  
CC EMBL; S52173; AAB24821.1; JOINED.  
CC EMBL; S52174; AAB24821.1; JOINED.  
CC EMBL; S52180; AAB24821.1; JOINED.  
CC EMBL; S52181; AAB24821.1; JOINED.  
CC EMBL; S52184; AAB24821.1; JOINED.  
CC EMBL; S52189; AAB24821.1; JOINED.  
CC EMBL; S52191; AAB24821.1; JOINED.  
CC EMBL; S52192; AAB24821.1; JOINED.  
CC EMBL; S52203; AAB24821.1; JOINED.  
CC EMBL; S52212; AAB24821.1; JOINED.  
CC EMBL; S52213; AAB24821.1; JOINED.  
CC EMBL; S52216; AAB24821.1; JOINED.  
CC EMBL; S52219; AAB24821.1; JOINED.  
CC EMBL; S52220; AAB24821.1; JOINED.  
CC EMBL; S52221; AAB24821.1; JOINED.  
CC EMBL; S52222; AAB24821.1; JOINED.  
CC EMBL; S52226; AAB24821.1; JOINED.  
CC EMBL; M76724; AAA58410.1; -  
CC EMBL; M84777; AAA51960.1; -  
CC PIR; A31108; RWHULB.  
CC PDB; 1ABX; 17-JUN-98.  
CC PDB; 1BHO; 18-NOV-98.  
CC PDB; 1BHQ; 18-NOV-98.  
CC PDB; 1IDN; 25-NOV-98.  
CC PDB; 1IDO; 01-AUG-96.  
CC PDB; 1JLM; 11-JAN-97.  
CC PDB; 1MLU; 07-AUG-02.  
CC Genew; HGNC:6149; ITGAM.  
CC MIM; 120980; -  
CC GO; GO:0008305; C:integrin complex; TAS.  
CC GO; GO:0004895; P:cell adhesion receptor activity; TAS.  
CC GO; GO:0007155; P:cell adhesion; TAS.  
CC InterPro; IPR000413; Integrin\_alpha.  
CC InterPro; IPR02035; VWFA.  
CC Pfam; PF01839; FG-GAP; 3.  
CC Pfam; PF00357; Integrin\_A; 1.  
CC Pfam; PF00092; vwa; 1.  
CC PRINTS; PR01185; INTEGRINA.  
CC PRINTS; PR00453; VWFADOMAIN.  
CC SMART; SM00191; Int\_alpha; 4.  
CC SMART; SM00327; VWFA; 1.  
CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
CC PROSITE; PS0234; VWFA; 1.  
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
CC Signal; 3D-structure; Repeat; Magnesium; Calcium.  
CC SIGNAL 1 16  
CC CHAIN 17 1152 INTEGRIN ALPHA-M.  
CC DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 1105 1128 POTENTIAL.  
CC DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).  
CC

QY	1021	FFGQIEEFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQAFVRSQTET	1080
Db	1036	FFGQIEEFNATLKGNSLFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQAFVRSQTET	1095
QY	1081	KVEFEVFNPLPLVGVSSVGLLILALITAAALYKLGPFKQYKDMSEGGPPGABPQ	1137
Db	1096	KVEFEVFNPLPLVGVSSVGLLILALITAAALYKLGPFKQYKDMSEGGPPGABPQ	1152
RESULT 2			
ITAM_MOUSE			
ID	ITAM_MOUSE	STANDARD; PRT; 1153 AA.	
AC	P0555;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-FEB-1991 (Rel. 17, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).		
DE	ITGAM.		
GN	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=88312584; PubMed=3044779;		
RA	Pyela R.;		
RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";		
RL	EMBO J. 7:1371-1378 (1988).		
RN	[2]		
RP	SEQUENCE OF 11-45 FROM N.A.		
RC	STRAIN=BALE/c; TISSUE=Spleen;		
RX	MEDLINE=86287312; PubMed=2942940;		
RA	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,		
RA	Larson R.S., Roberts T.M., Springer T.A.;		
RT	"A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648 (1986).		
RN	[3]		
RP	SEQUENCE OF 17-28.		
RX	MEDLINE=85188276; PubMed=3887182;		
RA	Springer T.A., Teplow D.B., Dreyer W.J.;		
RT	"Sequence homology of the LPA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";		
RL	Nature 314:540-542 (1985).		
CC	-1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO A IMPAIRED DEGRADATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.		
CC	-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.		
CC	-1- SUBCELLULAR LOCATION: Type 1 membrane protein.		
CC	-1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.		
CC	-1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.		
CC	-1- SIMILARITY: Belongs to the integrin alpha chain family.		
CC	-1- SIMILARITY: Contains 1 VWFA domain.		
CC	-1- SIMILARITY: Contains 7 FG-GAP repeats.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		

QY	1	PNLDTENAMTPOENARGFQOSVVOGSHVVVGAPOEIVAAHQRSLSLYOCDYSTGSCPEI	60
Db	17	PNLDTENAMTPOENARGFQOSVVOGSHVVVGAPOEIVAAHQRSLSLYOCDYSTGSCPEI	76
QY	61	RLOQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLPFGSNLRQOQOK	120
Db	77	RLOQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLPFGSNLRQOQOK	136
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPDFRMRKBLVSTIMBQKSKTLFSLMYSEEF	180
Db	137	FPEALRGCPQEDSDIAFLIDGSGSIIPDFRMRKBLVSTIMBQKSKTLFSLMYSEEF	196
QY	181	RTHFTKPEKONPNRSLKPTTOLLGRTHATGLRKVRELFNITNGARKNAKILFLJ	240
Db	197	RTHFTKPEKONPNRSLKPTTOLLGRTHATGLRKVRELFNITNGARKNAKILVVI	256
QY	241	TGCEKFGDPLGYEDVPIEDREGVIRYVLGFDAPRSEKSRQELNATVASKPPREDHVFQAN	300
Db	257	TGCEKFGDPLGYEDVPIEDREGVIRYVLGFDAPRSEKSRQELNATVASKPPREDHVFQAN	316
QY	301	NFEALKTQVQOLREKIPATGCTGTGSSSFHEMSQSGESAITSNGELISTVGSYDAG	360
Db	317	NFEALKTQVQOLREKIPATGCTGTGSSSFHEMSQSGESAITSNGELISTVGSYDAG	376
QY	361	GVFLYTSKEKSTFINTRVDSQMDAYLGVAAILLRNVQSLVGLGAPYQHIGLVAMPR	420
Db	377	GVFLYTSKEKSTFINTRVDSQMDAYLGVAAILLRNVQSLVGLGAPYQHIGLVAMPR	436
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIIGAPHYETRGQGVSVCP	480
Db	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIIGAPHYETRGQGVSVCP	496
QY	481	PRQORARWQCDVLYGECQGPGRFGAALTGLDVNGDKLTDVAGAPDEENRGAVILF	540
Db	497	PRG-RARWQCDVLYGECQGPGRFGAALTGLDVNGDKLTDVAGAPDEENRGAVILF	555
QY	541	HGTSGSGISPSHRSQRIAGSKLSPRLQYFGQSGGQDLTMDGLVLTVGAQGHVLLRSQ	600
Db	556	HGTSGSGISPSHRSQRIAGSKLSPRLQYFGQSGGQDLTMDGLVLTVGAQGHVLLRSQ	615
QY	601	PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT	660
Db	616	PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT	675
QY	661	YDLALDSGRPHSAVFNETKNSRRTQVLGTLQTCETLKLQLPNCIEDPVSPIVLRLNF	720
Db	676	YDLALDSGRPHSAVFNETKNSRRTQVLGTLQTCETLKLQLPNCIEDPVSPIVLRLNF	735
QY	721	SLVGTPLSAFGNLRPVLAEDAQLFTALPPFKNCGNDNICQDDLSTITPSFMSLDCLVVG	780
Db	736	SLVGTPLSAFGNLRPVLAEDAQLFTALPPFKNCGNDNICQDDLSTITPSFMSLDCLVVG	795
QY	781	GRPEFNVTVVRNDGDSYRTQVTFPPDLDSVRKYSTLQNSORSQSWELACESASTEV	840
Db	796	GRPEFNVTVVRNDGDSYRTQVTFPPDLDSVRKYSTLQNSORSQSWELACESASTEV	855
QY	841	SGALKSTCSINHPITPENSEVFNITFDVDSKASLGKLLKANKYTSNNMPRTNKTEF	900
Db	856	SGALKSTCSINHPITPENSEVFNITFDVDSKASLGKLLKANKYTSNNMPRTNKTEF	915
QY	901	QLELPVKYAVVWVTSHGVSQKYNLTASNTSRVMQHOYQVSNLQORSLSPLSLFLVLPV	960
Db	916	QLELPVKYAVVWVTSHGVSQKYNLTASNTSRVMQHOYQVSNLQORSLSPLSLFLVLPV	975
QY	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIQCDIP	1020
Db	976	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCORIQCDIP	1035

Query Match 99.1%; Score 5823.5; DB 1; Length 1152;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1123; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL; X07640; CAA30479.1; -;  
 DR EMBL; M14293; AAA39484.1; -;  
 DR PIR; S00551; S00551.  
 DR HSP; P11215; IABX.  
 DR XGD; MGI:96607; Itgam.  
 DR InterPro; IPR00413; Integrin\_alpha.  
 DR InterPro; IPR02035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Signal; Calcium; Repeat.  
 FT SIGNAL 1 16  
 FT CHAIN 17 1153 INTEGRIN ALPHA-M.  
 FT DOMAIN 17 1105 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1106 1129 POTENTIAL.  
 FT DOMAIN 1130 1153 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 31 84 FG-GAP 1.  
 FT REPEAT ? ? FG-GAP 2.  
 FT DOMAIN 164 350 VWA.  
 FT REPEAT 337 400 FG-GAP 3.  
 FT REPEAT 401 452 FG-GAP 4.  
 FT REPEAT 454 515 FG-GAP 5.  
 FT REPEAT 517 575 FG-GAP 6.  
 FT REPEAT 580 632 FG-GAP 7.  
 FT CA\_BIND 465 473 POTENTIAL.  
 FT CA\_BIND 529 537 POTENTIAL.  
 FT CA\_BIND 592 600 POTENTIAL.  
 FT SITE 1132 1136 GPFKR MOTIF.  
 FT DISULFID 66 73 BY SIMILARITY.  
 FT DISULFID 105 123 BY SIMILARITY.  
 FT DISULFID 654 711 BY SIMILARITY.  
 FT DISULFID 770 776 BY SIMILARITY.  
 FT DISULFID 999 1023 BY SIMILARITY.  
 FT DISULFID 1028 1033 BY SIMILARITY.  
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 696 696 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 734 734 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 772 772 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 801 801 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 881 881 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 907 907 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 941 941 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 994 994 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 1153 AA; 127480 MW; 178DB98BAECB0343 CRC64;

Query Match 75.6%; Score 4447; DB 1; Length 1153;  
 Best Local Similarity 73.5%; Pred. No. 1.5e-292;  
 Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

QY 1 FNLDTENAMTPOBNAEGFGQSVVLGGSRVVCAPQEIIVAAQNRGLYQCDYSTGSCBPI 60  
 DB 17 FNLDTEHPMTPOBNAEGFGQSVVLGGSTVVAAPQEAQAVNQTAGLYQCDYSTGRCHPI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTVVKGLCLFGLSGNLROQPOK 120  
 DB 77 FLOVPEAVNMSLGLSLAVTFPOLLACGPTVHONCKENTVYVGLCYLFGSNLLRPQQ 136  
 QY 121 PPEALRGCPQSDSDIAFLIDSGSII PHDFFRMKELNSTIMEOLKSKXTLSLMQYSBFB 180  
 DB 137 PPEALRGCPQESDIIVFLIDSGSINNIDFQMKEFVSTVMEQFKSKTSLFSLMQYSBFB 196  
 QY 181 RIHFTFKFONNPNRSLIKPTTOLLGRTHATGLRKVVVRLEFNITGNARKNAKILFLL 240  
 DB 197 RIHFTFNDFKPNFSPRSHVSPDKQLNGRTKTASGIRKVVVRLEFHTNARENAAILVVI 256  
 QY 241 TDGEKFGDPLGYEDVIPELDREGVIRYVLPDGFAPRSEKSRQELMTVASKPRPDHVFQAN 300  
 DB 257 TDGEKFGDPLDYKDVIPEDRAGVIRYVIGVGNFKNPKQSRRELDVTIASKPAGEHVFQVD 316  
 QY 301 NPEALKTQONOLREKI PAIEGTQTGSSSSFEHMSQBGFSAAITSGNPLLSVTGSDVWAG 360  
 DB 317 NPEALNTIQOQEKI PAIEGTQTGSSFEHMSQBGFSASITSGNPLLSVTGSDVWAG 376  
 QY 361 GVFLYTSXEKSTFINMTRVDSMDNDAYLGAAAIILRNVRQSLVGLGAPYQHIGLVAMFR 420  
 DB 377 GAFLYTSXDKVTFINMTRVDSMDNDAYLGAAVILRNVRQSLVGLGAPYQHIGLVAMFR 436  
 QY 421 QNTGWESNANVKTQICAYEGASILCSVDVDSNSTDLVILGAPHYEOTRGQGVSCPL 480  
 DB 437 ENFGTWEPTSIKSGISQISYFGASILCSVDMDADGNTNLILGAPHYEOTRGQGVSCPL 496  
 QY 481 PRGQARMQCDVILYGEQGPWGRFGAALTVLGVGVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
 DB 497 PRG-RARWQCEALLHGDQGHFAGHFGAALTVLGVGVNGDKLTDVAIGAPGEENQCAVYIF 555  
 QY 541 HGTSGGSLSPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
 DB 556 YGASTASLSASHSHRIIGAHPFSGLYQYFGQSLGSGKDLTMDGLMDLAVGAQCHLLLSAQ 615  
 QY 601 PVLRVKAIEMFNPREVARNVPCNDQVVKGEABEVRLHVQKSTRDLREGQIQSVVT 660  
 DB 616 PVLRLAETWEPSPKVAIVSACQEVLEKNDAGEVRVCLVRKNTKDLREGDIQSVTV 675  
 QY 661 YDLALDSGRPHSRVFNKTSRQTOVLGLTQCTETLKLQPNCEIDDPSPVILRLNF 720  
 DB 676 YDLALDPVRSIRAFDFDKNTTRRTQVFLMKQKCEETLKLILPDCVDSVSPILRLNY 735  
 QY 721 SLVGTPLSAFLNLRPVLAEDAQLFTALFPPEKKGNDNICODDISITFFSFWSLDCLVVG 780  
 DB 736 TLVGEPLSFGLNLRPVLAEDAQLFTALFPPEKKGNDNICODDISITFSWAMGLDVLVG 795  
 QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRL-ACRSASSTE 839  
 DB 796 GPQDFNMSVTLRNDGEDSYGTQVTVYVPSGLSYRKDSASQNPFTKPKWFKVPAIRSSSE 855  
 QY 840 VSGALKSTSCSINHPI PPESEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTE 899  
 DB 856 GHGALKSTTWNINHPI PPNSEVTFNITFDVDSHAGFGNKLKLLKAVASENNMERTHTK 915  
 QY 900 FOLELPVKYAVIMVTSHGVSTKYNLFTASENTRVMQHOYQVSNLQORSPLISLVPLVP 959  
 DB 916 FOLELPVKYAVIMVTSESSIRYLNFTASEMTSKVIHQYQFNGLGORSPLSPSVVFWIP 975  
 QY 960 VRLNQVTIWDPEPOTFSENSTCHTKERLPSHSDLAELERKAPVWNCISIAVCORICDI 1019  
 DB 976 VQINNVTVWDHPQVIFQNSLSAACHTKQKSPPHSNFRDQLERTPVLCNSVAVCKRIQCDL 1035  
 QY 1020 PFFGIGQEEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDVSTFTLLPGQCAFVRSQTE 1079  
 DB 1036 PSENTQIEFNVTLKNLSFDWYIKTSHGHLILVSSSTELFNDSAFALLPGQESVYRSKTE 1095  
 QY 1080 TKVEPPEFVNPDLPLIVSSVCGHLLILITAAALKGLGFEKQYKDMSEGGCPGAPQ 1137  
 DB 1096 TKVPEFVHNVPPLIVSSIGGLVLLALITAGLYKGLGFEKQYKDMNEAAPQDAPPQ 1153

RESULT 3  
ITAX HUMAN  
ID ITAX HUMAN STANDARD; PRT; 1163 AA.  
AC P20702;  
DT 01-FEB-1991 (Rel. 17, Last Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95  
DE alpha chain) (leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).  
GN ITGAX OR CD11c.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=88166645; PubMed=3327687;  
RA Corbi A.L.; Miller L.J.; O'Connor K.; Larson R.S.; Springer T.A.;  
RT "CDNA cloning and complete primary structure of the alpha subunit of  
RT a leukocyte adhesion glycoprotein, p150,95";  
RL EMBO J. 6:4023-4028(1987).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=90153906; PubMed=2303426;  
RA Corbi A.L.; Garcia-Aguilar J.; Springer T.A.;  
RT "Genomic structure of an integrin alpha subunit, the leukocyte  
RT p150,95 molecule.";  
RL J. Biol. Chem. 265:2782-2788(1990).  
RN [3]  
RN ERRATUM.  
RA Corbi A.L.; Garcia-Aguilar J.; Springer T.A.;  
RL J. Biol. Chem. 265:12750-12751(1990).  
RN [4]  
RN SEQUENCE OF 20-43.  
RX MEDLINE=87167596; PubMed=3549901;  
RA Miller L.J.; Wiebe M.; Springer T.A.;  
RT "Purification and alpha subunit N-terminal sequences of human Mac-1  
RT and p150,95 leukocyte adhesion proteins.";  
RL J. Immunol. 138:2381-2383(1987).  
CC [1]- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT  
CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL  
CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY  
CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.  
CC [2]- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X  
CC ASSOCIATES WITH BETA-2.  
CC [3]- SUBCELLULAR LOCATION: Type I membrane protein.  
CC [4]- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND  
CC GRANULOCYTES.  
CC [5]- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC [6]- SIMILARITY: Belongs to the integrin alpha chain family.  
CC [7]- SIMILARITY: Contains 1 VWFA domain.  
CC [8]- SIMILARITY: Contains 7 FG-GAP repeats.  
CC [9]- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M81695; AAA51620.1; -;  
CC EMBL; Y00093; CAA68283.1; -;  
CC EMBL; M29145; -; NOT ANNOTATED CDS.  
CC EMBL; M29487; AAA51620.1; ALT\_SEQ.  
CC EMBL; M29482; AAA51620.1; JOINED.  
CC EMBL; M29483; AAA51620.1; JOINED.  
CC EMBL; M29484; AAA51620.1; JOINED.  
CC EMBL; M29485; AAA51620.1; JOINED.

DR EMBL; M29486; AAA51620.1; JOINED.  
DR PIR; A36594; RWHUIC.  
DR PDB; 1N3Y; 18-FEB-03.  
DR Genew; HGNC:6152; ITGAX.  
DR MIM; 151510; -;  
DR GO; GO:0008305; C:integrin complex; TAS.  
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
DR GO; GO:0004872; F:receptor activity; TAS.  
DR GO; GO:0007155; P:cell adhesion; TAS.  
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWFA.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PRO1185; INTEGRINA.  
DR PRINTS; PRO0453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWFA\_1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Magnesium; Calcium; Repeat; 3D-structure.  
FT SIGNAL 1 19 INTEGRIN ALPHA-X.  
FT CHAIN 20 1163 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 20 1107 POTENTIAL.  
FT DOMAIN 1108 1128 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 34 87 FG-GAP 1.  
FT REPEAT ? ? FG-GAP 2.  
FT DOMAIN 165 351 VWFA.  
FT REPEAT ? ? FG-GAP 3.  
FT REPEAT 402 453 FG-GAP 4.  
FT REPEAT 455 517 FG-GAP 5.  
FT REPEAT 518 576 FG-GAP 6.  
FT REPEAT 581 633 FG-GAP 7.  
FT CA\_BIND 466 474 POTENTIAL.  
FT CA\_BIND 530 538 POTENTIAL.  
FT CA\_BIND 593 601 POTENTIAL.  
FT SITE 1131 1135 GPEKE MOTIF.  
FT DISULFID 69 76 BY SIMILARITY.  
FT DISULFID 108 126 BY SIMILARITY.  
FT DISULFID 655 712 BY SIMILARITY.  
FT DISULFID 771 777 BY SIMILARITY.  
FT DISULFID 848 863 BY SIMILARITY.  
FT DISULFID 998 1022 BY SIMILARITY.  
FT DISULFID 1027 1032 BY SIMILARITY.  
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 735 735 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 899 899 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 490 490 G -> A (IN REF. 2).  
FT CONFLICT 756 756 L -> D (IN REF. 2).  
SQ SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3F62A473 CRC64;  
Query Match 58.6%; Score 3446; DB 1; Length 1163;  
Best Local Similarity 60.5%; Pred No. 8e-225;  
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;  
QY 1 FNLDTENAMTQENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLIQDYSTGSCPEI 60  
DB 20 FNLDTEELTAFRVDSAGFGDSVQVYANSVWVVGAPQKITAAQNTGGLYCCGYSTGACPEI 79  
QY 61 RLQVPVEAVNMSLGLSLAATTPBQLACGPTVHOTCSNTVYVXGLCLFLFGLNLRQQPK 120  
DB 80 GLQVPPPEAVNMSLGLSLASTTSPQLACGPTVHCCGNMYLTGLCLLFT--QLTQR 137  
QY 121 FPEALRCQPBDSIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSSEP 180

Db 138 LPVSRQECPRQEDIVFLIDSGSISRNFMATMNFVRAVISQPRSTQFSLSQFSNKF 197  
QY 181 RIHTFKEFQNNPNRSLRIKFIITQLGRTHITATGLRVVRELFNITMGARKNAFKILFL 240  
Db 198 QTHFTFEFRSTNPLSLLASVHLOQFTVATATQVNVHRLPHASYGARRDATKILIVI 257  
QY 241 TDGKFKGDPLGYEDVIVPELREGVIRVILVFGDAPRSEKSRQELNTVASPPDRHVPQAN 300  
Db 258 TDGKKGSDLDYKDVIPMAAAGIIRYAIGVGLAFQNRNSWKLNDIASPESQEHIPKVE 317  
QY 301 NFEALTKVQNLREKIFAIBGTQTGSSSSFEHMSQGFSAAITNSGPLLSTVGSVDWAG 360  
Db 318 DFDALKDIQNLKEKIFAIBGTETISSSFELEMAQGFSAVFTPDGVLGAVGSPMSG 377  
QY 361 GVPLYTSEKSTPTNMTVSDNDAYLGYAAAILLRVQSLVGLGAPRYOHLGLVAMER 420  
Db 378 GAFLYPPNMSGPTFINMSQENVNDSYLSYSLALWKGVSILVGLGAPRYOHTGKAVIFT 437  
QY 421 QNTGMESNANVKTQIGAFYGASLCSVDVDSNGSTDVLVIGAPHYVYQTRGGQVSCPL 480  
Db 438 QVSRQWRKAEVGTQIGSYFGSLCSVDVDTGSTDVLVIGAPHYVYQTRGGQVSCPL 497  
QY 481 PRQARWQCDVLYGQGGQFWRGFGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540  
Db 498 PRGWR-RWVCDVLYGQGGQFWRGFGAALTIVLGVNGDKLTDVAIGAPBEENRGAVYLF 556  
QY 541 HGTSGSGISFSHQRJAGSKLSPLOYFGOSLSGGDLTMDGLVDTVGAQGHVLLRSQ 600  
Db 557 HGVLFPSISFSHQRJAGSKLSPLOYFGOSLSGGDLTMDGLVDTVGAQGHVLLRSQ 616  
QY 601 PVLRVKAIMEFNPREVARNVFCNDQVYKGEAGEVRVCLHVOKXSTRDLRREGIQSVVT 660  
Db 617 PVLAVGVSMQFIPAEIPRSAFEQVSVSQTUVQSNICLYIDKRSKNLIGSRDLQSSVT 676  
QY 661 YDLALDSGRPHSAVENETKNSRTOVLGLTQCTETKLPCLPCLEQVPSVILRLNF 720  
Db 677 LDIALDPGLSPRATQETKNSRVRVLGKAHCENFLLPSCVEDSVPTITLRLNF 736  
QY 721 SLVGTPLSAFNGRLPVLAEADQRLTALFPFERNCGNDNICDDLSITFSFMSLDCLVVG 780  
Db 737 TLVGKPLLAFLNRLPMLAAQRYFTASLPFERNCGADHICQDNLGSPFPLKSLVVG 796  
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSLACASSTEV 840  
Db 797 SNLELNAEVMVNDGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWSLACASSTEV 854  
QY 841 SGALKSTSCSINPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 900  
Db 855 SOGTWSTSCRINELIFRGGAQITFLATFDSVPAVLGDRLLTANVSSENNPRTSKTFF 914  
QY 901 QLELPVKYAYVTVSSHEQTKYLNFTAS-ENTSRVMOHQYQVNSLQSRSLPISLVFLVP 959  
Db 915 QLELPVKYAYVTVSSHEQTKYLNFTAS-ENTSRVMOHQYQVNSLQSRSLPISLVFLVP 974  
QY 960 VRLNQTVIMDRPOVTFSENLSSYCHTKERLPSHSDFLAELRKPVNVCSIAVOCIOCDI 1019  
Db 975 VELNQAEVMDVESHVPQNPSPSCSEKTAAPASDFLAHOKNPVLDCSIAGLAPRCVD 1034  
QY 1020 PFGIOBEFNATKGNLSDFWYIKTSHNLLIVSTAEILFNDSVFTLLPQGAQFVRSQTE 1079  
Db 1035 PFSVOEELDTLKGHLSFGWRQILQKXVSVVVAEITFDTSVSOLPQGAQFVRSQTE 1094  
QY 1080 TKVEPEVFNPLIIVGSSVGGLLALITAAVLYKGFPRKQYKDMWSE 1128  
Db 1095 TVLEKVKVNPPLIIVGSSVGGLLALITAAVLYKGFPRKQYKDMWSE 1143

## RESULT 4

ID ITAD\_HUMAN  
AC Q13349: Q15575; Q15576;  
BT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).  
GN ITGAD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=96111956; PubMed=8777714;  
RA Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,  
RA Staunton D.E., Gallatin W.M.;  
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-  
RT 3.";  
RL Immunity 3:683-690 (1995).  
RN [2]  
RP SEQUENCE OF 1-235 FROM N.A.  
RX MEDLINE=20187620; PubMed=10722744;  
RA Noti J.D., Johnson A.K., Dillon J.D.;  
RT "Structural and functional characterization of the leukocyte integrin  
RT gene CD11d. Essential role of Sp1 and Sp3.";  
RL J. Biol. Chem. 275:8959-8969 (2000).  
RN [3]  
RP SEQUENCE OF 457-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.  
RX MEDLINE=96257236; PubMed=8666289;  
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;  
RT "Cloning and chromosomal localization of a novel gene encoding a human  
RT beta 2-integrin alpha subunit.";  
RL Gene 171:291-294 (1996).  
RN [4]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99059842; PubMed=9841932;  
RA Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,  
RA Hoffman P.A., Staunton D.E., Sochner B.S.;  
RT "alpha2beta2 integrin is expressed on human eosinophils and functions  
RT as an alternative ligand for vascular cell adhesion molecule 1  
RT (VCAM-1).";  
RL J. Exp. Med. 188:2187-2191 (1998).  
RN [5]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99370002; PubMed=10438935;  
RA Van der Vieren M., Crowe D.F., Hoekstra D., Vazeux R., Hoffman P.A.,  
RA Grayson M.H., Sochner B.S., Gallatin W.M., Staunton D.E.;  
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
RT binding interface between I domain and VCAM-1.";  
RL J. Immunol. 163:1984-1990 (1999).  
CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND  
CC VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS  
CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-  
CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENESCENT ERYTHROCYTES  
CC FROM THE BLOOD.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D  
CC ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL  
CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON  
CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN  
CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWF A DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWF A domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).  
CC -----

DR EMBL; U37028; AAB38547.1; -;  
DR EMBL; U40274; AAB60634.1; -;  
DR EMBL; U40275; AAB60635.1; -;  
DR EMBL; U40276; AAB60636.1; -;  
DR EMBL; U40277; AAB60637.1; -;  
DR EMBL; U40279; AAB60638.1; -;  
DR EMBL; U40278; AAB60638.1; JOINED.  
DR EMBL; AF187881; AAF62875.1; -;  
DR HSSP; P11215; IABX.  
DR Genew; HGNC:6146; ITGD.  
DR MIM; 602453; -;  
DR GO; GO:0008305; C: integrin complex; TAS.  
DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.  
DR GO; GO:0016337; P: cell-cell adhesion; NAS.  
DR GO; GO:0007160; P: cell-matrix adhesion; NAS.  
DR GO; GO:0006955; P: immune response; NAS.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWA\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PRO1185; INTEGRINA.  
DR PRINTS; PRO0453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 4.  
DR SMART; SM00327; VWA\_1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Repeat; Calcium;  
KW Magnesium.  
FT SIGNAL 1 17 POTENTIAL.  
FT CHAIN 18 1162 INTEGRIN ALPHA-D.  
FT DOMAIN 18 1100 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1101 1124 POTENTIAL.  
FT CYTOPLASMIC 1125 1162 (POTENTIAL).  
FT REPEAT 32 85 FG-GAP 1.  
FT REPEAT 33 86 FG-GAP 2.  
FT REPEAT 34 87 VWA.  
FT REPEAT 35 88 FG-GAP 3.  
FT REPEAT 36 89 FG-GAP 4.  
FT REPEAT 40 92 FG-GAP 5.  
FT REPEAT 45 96 FG-GAP 6.  
FT REPEAT 51 102 FG-GAP 7.  
FT REPEAT 58 109 POTENTIAL.  
FT CA\_BIND 465 473 POTENTIAL.  
FT CA\_BIND 530 538 POTENTIAL.  
FT CA\_BIND 593 601 POTENTIAL.  
FT SITE 1127 1131 GFGR MOTIF.  
FT DISULFID 67 74 BY SIMILARITY.  
FT DISULFID 106 124 BY SIMILARITY.  
FT DISULFID 655 710 BY SIMILARITY.  
FT DISULFID 769 775 BY SIMILARITY.  
FT DISULFID 846 861 BY SIMILARITY.  
FT DISULFID 994 1018 BY SIMILARITY.  
FT DISULFID 1023 1028 BY SIMILARITY.  
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 691 691 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 873 873 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 957 957 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1046 1046 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 500 500 MISSING (IN REF. 2).  
FT CONFLICT 515 518 GHFW -> ATP (IN REF. 2).  
FT CONFLICT 825 825 L -> V (IN REF. 2).  
FT CONFLICT 984 984 V -> A (IN REF. 2).  
FT SEQUENCE 1162 AA; 126885 MW; F296AIA3545D77D CRC64;

Query Match 57.6%; Score 3388; DB 1; Length 1162;  
Best Local Similarity 58.9%; Pred. No. 6.7e-221;  
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFOENARGFGQSVVOLQSGRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCSEPI 60  
DB 18 FNLQVEBETIFQEDAGFGQSVVQFGSRLVVGAPLEVVAAANTGRLVCAAAATGWCQPI 77  
QY 61 RLQVPVAVNNMGLSLAAATTSPPOLLAAGTPTVQTCSENTYVVKGLCFPLFGSNLRQQQPK 120  
DB 78 PLHIRPEAVNNMGLSLTAAATNGSRLACGPTLHERVCCENSYSKSGSCULLSRW-EITQT 136  
QY 121 FPEALRGCPQSDIAFLIDGSGSIIIPHDPRMKELVSTIMEQKKKSKTLFSLMOYSBEF 180  
DB 137 VDATPECPHOEMDIVFLIDGSGSIDQDNFQMGKGFQVAVMGQEGTDTLFAIMOYSNLL 196  
QY 181 RIHTFTEFQNNPNRSLIKPITQLGRHTHTATGLRKVVRELPMITNGARKNAKIFILL 240  
DB 197 KIHFTFTQFRTSPSQSLVDPVQLKGLTFTATGILTVTQLPHHKGARKSARKILLVI 256  
QY 241 TDGKFGDPLGYEDVIPLEDEGVIRYVLGPDGAFRSKSRQELNTVASKEPRDHVFOAN 300  
DB 257 TDGQYKDPLEYSVIPAOKAGIIRIYAGVGHAFQGPATROELNTISSAPQDHYVFKVD 316  
QY 301 NFEALKTVMQNLREKIPAEIGTQTGSSSPHEHMSQEGFSAATTSNGELLSTVSGSYDWAG 360  
DB 317 NFAALGSIQKQLEKIVAVEGTQSRASSFOHEMSQEGFSTALTMGDLFLGAVGFSFMSG 376  
QY 361 GVFLYTSKEKSTFTNMTFVDSMDNDAYLGYAAAIIILNRVQSLVGLGAPRYOHIGLVANFR 420  
DB 377 GAPLYPPNMSPTFINMSQENVMDSDSYLGSTELALMKGQNLVLGAPRYOHTGKAVIT 436  
QY 421 QNTGHWESNANVKTQIGAYFGASLCSDVDNSGSTDVLVIGAPRYEYQTRGGQSVSCPL 480  
DB 437 QVSRQWRKKAETGTQIGSYFGASLCSDVDNSDGTDLILGAPRYEYQTRGGQSVSCPL 496  
QY 481 PRGORARWQCDVLYGQGPWGRFGAALTVDGVNGDKLTDVAIGAPCEEDNRCAYLF 540  
DB 497 PRGORVQWQCDVLYGQGPWGRFGAALTVDGVNGDKLTDVAIGAPCEENRCAYLF 556  
QY 541 HGTSGSGISPSHSORLAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600  
DB 557 HGASESGISPSHSORLAGSKLSPLQYFGQALSGQDLTQDGLANDLAVGARGQVLLLRSL 616  
QY 601 PVLNVKAIMFNPREVARNVFCNDQVVKGEAGEVRLHVQKSTRDRLEGQIQSVVT 660  
DB 617 PVLKVGVMARFSPVEVAKAVRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSVR 674  
QY 661 YDLALDSGRPHSRVFNETKNSRROTQVLGLTCTCETKLQLPNCIEDPSPIVLRNLF 720  
DB 675 FDALDPRLTSRAIFNETKNPILTRKTLGLGHCHETKLLLPDCVEDVVSPIILHLNF 734  
QY 721 SLVGTPLSAFQNLRLVLAEDAQRLLFTALFPPEKNGNDNICODDLSTITFSFMSLDCLVVG 780  
DB 735 SLVREPIFSPONLRPLVAVGSDLTFTASLPPEKNGQDGLCEGLGVTLSPSLGLTLTVG 794  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPLDLVSEKVTSLQNRORSRSLACSSASTEV 840  
DB 795 SSLEINLVTVWNAGEDSYRTVVSLLYPAGLUSHRVSGAQKPPHQSURLACETV-PTED 853  
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTKTEF 900  
DB 854 EG-LRSSRCSVNHPIFHEGSGNGTFTVTFDVSYKATLGORMLMRASSENKASSSKATF 912  
QY 901 QLELPVKVAVMVTSHGVSTKYNF-TASENTRVMQHQVQVSNLQORSIPISLVPLVP 959  
DB 913 QLELPVKVAVMVTSHGVSTKYNF-TASENTRVMQHQVQVSNLQORSIPISLVPLVP 972  
QY 960 VRLNQTVMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVWNCSTAVCQRIQCDI 1019  
DB 973 VLVINGVAVWVMEAPSQSL--PCVSEKPPQSHDFLTQISRSFPLDCSTADCLQFRCDV 1030  
QY 1020 PFGIQSEFNATLKNLGFQWYIKTSNNHLLIVSTAEILFNDSVFTLLPGQCAFVRSQTE 1079  
DB 1031 PSFSVQBELDFTLKNLSFGVWRBTLQKVLVWVSAEITDTSVYSQIPQGEAFMRQME 1090  
QY 1080 TKVEPPEVNPPLPLIVGSSVGLLILLALITAAALYKLGFFKRYKDMWSE 1128



```

FT CARBOHYD 65 65 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 89 89 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 188 188 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 649 649 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 670 670 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 726 726 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 730 730 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 862 862 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 885 885 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 897 897 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1060 1060 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1071 1071 N-UNLINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC 954 954 Q -> QGVHGLVEMQTSKQILCEPAGDAHTVGAQGEPLCP
PWGVSAPRDNIRAGPCR (in isoform 2).
FT FTIG=VSP 002738.
R -> W (IN REF. 1 AND 2).
Y -> I (IN REF. 2).

FT CONFLICT 214 214
FT CONFLICT 660 660
FT STRAND 155 162
FT STRAND 164 164
FT TURN 165 165
FT HELIX 169 185
FT TURN 186 188
FT STRAND 191 198
FT STRAND 202 206
FT HELIX 208 213
FT HELIX 217 221
FT TURN 222 223
FT STRAND 229 229
FT HELIX 233 243
FT TURN 244 244
FT HELIX 247 249
FT TURN 250 250
FT TURN 253 254
FT STRAND 256 263
FT HELIX 274 276
FT TURN 277 278
FT STRAND 280 286
FT HELIX 288 290
FT TURN 293 297
FT TURN 298 299
FT HELIX 300 302
FT HELIX 307 310
FT STRAND 311 314
FT TURN 318 318
FT HELIX 319 328
FT TURN 329 330
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match 26.1%; Score 1536.5; DB 1; Length 1170;
Best Local Similarity 34.2%; Pred. No. 1.2e-95;
Matches 401; Conservative 209; Mismatches 463; Indels 101; Gaps 37;

QY 1 FNLDTENAMTFQ--ENARGFGQSVQVQGSRRVVVGAPOEIVAAQNGSLYQCQDYSYGSC 58
DB 26 YNLDVRCARSPFPFRAGRHFGYEVLVQV-GNGVIVGAPGE---GNSGTSLYQCQSGTGCL 81
QY 59 PTLQVPEAVNMSGLSLAATSPOLLACGTVTQTCSTNTYVGLCLPLFGSNLR--- 115
DB 82 PVTLR-GSNYTSKYLGMTLATDPTDGSILACDPLSRTCDQNTYLSGLCYLFRQNLQGP 140
QY 116 -QQPKPFPALRGCPQEDSDIAFLIDGSGSIIIPDFRRMKELVSTIMEOLKSKTLESLM 174
DB 141 LQRPQFPQECIKG----NVDLVFPGSVSLQDFQKILDFPKDVKKLSTSYQPAAY 196
QY 175 QYSEFRHIFTPEKFNQNNPRSLIKPIITOLLGRHTATGLRKVVRELFNITNGARKNAP 234
DB 197 QFSTSYKTEPFDGSDYVKRKPDPALLXVHMLLTNTFPAINIVATEVFRBELGAPDAT 256
QY 235 KILFLATDGEKFDPIGYEDVIELOREGVIRVILGFGDAFRSEKSRQINTVASKCPED 294
DB 257 KVLIIITDGS--ATDSGNIDAAND-----IIRVIIGIKGHFKQESQETILKFPASKPASE 309
QY 295 HVFQANNFEALTKTVQNLREKIPAEIGCTQTGSSSSFEHEMSQEGFAAITSNGPLISTV 354

```

## RESULT 6

## ITAL MOUSE

## ID ITAL MOUSE

## STANDARD;

## PRT; 1163 AA.

## AC P24063;

## DT 01-MAR-1992 (Rel. 21, Created)

## DT 01-MAR-1992 (Rel. 21, Last sequence update)

## DT 28-FEB-2003 (Rel. 41, Last annotation update)

## DE Integrin alpha-1 precursor (Leukocyte adhesion glycoprotein LFA-1

## DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)

## DE (CD11a).

## GN ITGAL OR LFA-1.

## OS Mus musculus (Mouse).

## OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91269576; PubMed=2051027;  
RA Kaufmann Y., Tseng B., Springer T.A.;  
RT "Cloning of the murine lymphocyte function-associated molecule-1  
alpha-subunit and its expression in COS cells.";  
RL J. Immunol. 147:369-374(1991).  
RN [2]  
RP SEQUENCE OF 24-42.  
RX MEDLINE=9518276; PubMed=3887182;  
RA Springer T.A., Teplov D.B., Dreyer W.J.;  
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion  
glycoproteins and unexpected relation to leukocyte interferon.";  
RL Nature 314:540-542(1985).  
CC -!- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,  
ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA  
INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL  
MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES  
AND MONOCYTES. NICE EXPRESSING A NULL MUTATION OF THE ALPHA-L  
SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED  
LEUKOCYTES RECRUITMENT.  
CC -!- SUBUNIT: HETERO DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L  
ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: LEUKOCYTES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSTR) IS A VWF A DOMAIN. INTEGRINS  
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWF A domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M60778; AAA39426.1; -.  
DR PIR: I56126; I56126.  
DR HSSP: P20701; 1LFA.  
DR MGD: MGI:96606; Itgal.  
DR InterPro: IPR000413; Integrin\_alpha.  
DR InterPro: IPR002035; VWF A.  
DR Pfam: P201839; FG-GAP; 3.  
DR Pfam: P200357; Integrin\_A; 1.  
DR Pfam: P200092; vwa; 1.  
DR PRINTS: PR01185; INTEGRINA.  
DR PRINTS: PR00453; VWPADOMAIN.  
DR SMART: SM00327; VWA; 1.  
DR SMART: SM00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE: PS00234; VWF A; 1.  
DR PROSITE: PS00242; INTEGRIN\_ALPHA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Magnesium; Calcium;  
KW Repeat.  
FT SIGNAL 1 23  
FT CHAIN 24 1163  
FT DOMAIN 24 1084  
FT TRANSMEM 1085 1108  
FT DOMAIN 1109 1163  
FT REPEAT 39 88  
FT REPEAT ? ?  
FT DOMAIN 148 334  
FT REPEAT ? ?  
FT REPEAT 399 454  
FT REPEAT 455 514  
FT REPEAT 516 573  
FT REPEAT 576 628  
FT CA\_BIND 466 474  
FT POTENTIAL.

FT	CA_BIND	528	536	POTENTIAL.
FT	CA_BIND	588	596	POTENTIAL.
FT	SITE	1111	1115	GFPR MOTIF.
FT	DISULFID	70	77	BY SIMILARITY.
FT	DISULFID	103	126	BY SIMILARITY.
FT	DISULFID	147	139	BY SIMILARITY.
FT	DISULFID	651	705	BY SIMILARITY.
FT	DISULFID	767	773	BY SIMILARITY.
FT	DISULFID	840	856	BY SIMILARITY.
FT	DISULFID	993	1009	BY SIMILARITY.
FT	DISULFID	1017	1048	BY SIMILARITY.
FT	CARBOHYD	85	86	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	185	185	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	270	270	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	668	668	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	696	696	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	724	724	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	775	775	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	880	880	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	890	890	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	899	899	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	927	927	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1056	1056	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	1163 AA;	128343 MW;	A7A30784898232F CRC64;

Query Match 25.8%; Score 1516.5; DB 1; Length 1163;  
Best Local Similarity 33.9%; Pred. No. 2.6e-94;  
Matches 396; Conservative 216; Mismatches 466; Indels 89; Gaps 36;

QY	1	ENLDTENAMTFQENA-RGFGQSVVOLQSGSRVVVGAPOBIVAAQNRGLVQCQDYSTGSCFP	59
DB	24	YNLDRTPQSFIAQGRHFGYQVLQEDG-VVGAPOB---GDTGGLTACTSTSSFCQF	79
QY	60	IRLQVEVAVNMGLSLAATSPPLLACGPTVHTQCSNTYVVKGLCLFOSNLQQPQ	119
DB	80	VSLH-GSNHTSKYLGMTLATDAAGSLACDGLSRTCDQNTYLSGLYLPQSLGPM	138
QY	120	KPEALRGCFQEDSIAFLIDGSGTIIPDPRMKELVSTIMEQLKSKTSLFSLMOYSEE	179
DB	139	QNRPAYQECMKGVLDVLPFDGSGSLDRKDFKILFEMCDVMRKLNTSYQFAAVGSTD	198
QY	180	FRHPTTFKEF-QNNPNRSLRPIPTQLLGRTHATGLRKVVRBELFNITNGARKAPKILF	238
DB	199	CRTEPTFLDYVQKNKFDVLLGSVQPMPLTNTFRATVVAHVFKBESGARPDAKVLV	258
QY	239	LLTDGKFP--GDPLGYEDVPELDREGVIRYVLGRGDAPRSEKSRQELNTVASKPRDRV	296
DB	259	IITDGEASDKNISAAHD-----ITRYIIGIKRHFVSQKQKTLHFASEPVEEFV	309
QY	297	FQANNPEALKTVQNLREKIFAIEGTQTGSSSSFFHEMSQEGFSAITNGPPLLSTVGSY	356
DB	310	KILDTPFKLQDLFTDLQRRIVAEQTNQDLTSFNWELSSSGISADLSKGVAVGVAGAK	369
QY	357	DWAGGVF-LYTSKEKSTFTNTRVDSMDNDAYLGAA-ILLRNVQSLVIGAPRYOHLG	414
DB	370	DWAGGFLLDREDLQGAFTVQGEPLTSDVRGGLTYVTWMTSRSRPRLAAGAPRYQHV	429
QY	415	LVAMPR--QNTGMWBSNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVIGAPHYEQTRG	472
DB	430	QVLLQAPAGGRWQTKLEGTQIGSVFGGLCSVDLDQDGEAEILLIGAPLFGQRG	489
QY	473	QGVSVCLPRGQARWQCDVLYGQGGQPMGRFGAALTVLGVNGDKLTDVIGAPGED	532
DB	490	GRVFTY--QRRQSLPEMVSELQGGPGYPLGRFGAALTALTDINGDLTDVAVGAPLE-	545
QY	533	NRGAVYLPHTSGSGISPSHSQRIAGSKLSRLQYFGOSLSGGQDLTMDGLVDLTVAQG	592
DB	546	-QGVYIFNGKFG-GLSPQPSORIQAQVFPGRIFGRSIHVVKDLGGDLADVVVAGS	603
QY	593	HVLLRSQPLRVKAKIMEFNPFEVARNFECDQNVVKGKEAG-SYRVCLHVQKSTRDLR	651

```

RESULT 7
ITAE_MOUSE
ID ITAE_MOUSE STANDARD; PRT; 1167 AA.
O60677;
AC AC
DT DT
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-E precursor (Integrin alpha M290).
GN IIGAE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
O1 _TaxID=10090;
[1]
RPN SEQUENCE FROM N.A.
RPR STRAIN=AKR;
RRC MEDLINE=95187992; PubMed=7882170;
RRA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,
RRA Kilshaw P.J., Weiss J.H.;
RRT "Murine M290 integrin expression modulated by mast cell activation.";
RRT Immunity 1,393-403(1994)
CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL
CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E
CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF
CC INTRA-EPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF
CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR
CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE
CC INTESTINAL WALL.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

```

CC	-1-	DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A WFPA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC	-1-	SIMILARITY: Belongs to the integrin alpha chain family.
CC	-1-	SIMILARITY: Contains 1 WFPA domain.
CC	-1-	SIMILARITY: Contains 7 FG-GAP repeats.
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	ENBL;	U12236; AAC52142.1; --
DR	HSPG;	P11215; IABX.
DR	MGD;	MGI:1298377; Itgae.
DR	InterPro;	IPR000413; Integrin_alpha.
DR	InterPro;	IPR002035; VWF_A.
DR	Pfam;	PF01839; FG-GAP_3.
DR	Pfam;	PF00357; integrin_A; 1.
DR	Pfam;	PF00052; vwa; 1.
DR	PRINTS;	PR01185; INTEGRINA.
DR	PRINTS;	PR00453; VWFADOMAIN.
DR	SMART;	SM00191; Int_alpha; 3.
DR	SMART;	SM00327; VWA; 1.
DR	PROSITE;	PS00242; INTEGRIN_ALPHA; 1.
DR	PROSITE;	PS50234; VWFA; 1.
KW	Integrin;	Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW	Signal;	Repeat; Magnesium;
KW	Calcium.	
FT	CHAIN	1 19 BY SIMILARITY.
FT	CHAIN	20 1167 INTEGRIN ALPHA-E.
FT	CHAIN	20 181 INTEGRIN ALPHA-E LIGHT CHAIN.
FT	CHAIN	183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.
FT	DOMAIN	20 1114 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1115 1137 POTENTIAL.
FT	DOMAIN	1138 1167 CYTOPLASMIC (POTENTIAL).
FT	REPEAT	? ? FG-GAP 1.
FT	REPEAT	? ? FG-GAP 2.
FT	DOMAIN	149 192 X-DOMAIN (EXTRA DOMAIN).
FT	DOMAIN	193 384 VWFA.
FT	REPEAT	? ? FG-GAP 3.
FT	REPEAT	449 501 FG-GAP 4.
FT	REPEAT	503 564 FG-GAP 5.
FT	REPEAT	566 631 FG-GAP 6.
FT	REPEAT	634 686 FG-GAP 7.
FT	CA_BIND	514 522 POTENTIAL.
FT	CA_BIND	578 586 POTENTIAL.
FT	CA_BIND	646 654 POTENTIAL.
FT	DOMAIN	185 191 GLU-RICH (ACIDIC).
FT	SITE	1140 1144 GFGR MOTIF.
FT	DISULFID	72 83 BY SIMILARITY.
FT	DISULFID	130 154 BY SIMILARITY.
FT	DISULFID	638 754 BY SIMILARITY.
FT	DISULFID	814 820 BY SIMILARITY.
FT	DISULFID	884 898 BY SIMILARITY.
FT	DISULFID	998 1023 BY SIMILARITY.
FT	DISULFID	1031 1047 BY SIMILARITY.
FT	CARBOHYD	51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	256 256 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	314 314 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	718 718 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	773 773 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	829 829 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	846 846 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	925 925 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	968 968 N-LINKED (GLCNAC. .) (POTENTIAL).

```
FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1066 1066 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 128993 MW; 83331C115DCCCFD CRC64;

Query Match
Best Local Similarity 28.6%; Pred. No. 1.3e-68;
Matches 353; Conservative 213; Mismatches 462; Indels 205; Gaps 43;

Qy 1 FNLDTEHA--MTPDENARGFGQSVVQLGGRRVVVGAPOEIVANQRGSG-----LYQCOY 52
Db 20 FNMVDNMAWTALQFGAPAVLSLHLDPSS-----NQTCLLVARRSSNNTAALYRCAL 74
Qy 53 STGCEBIRLQVPEAVNMSLGLSLAATT--SPQLLAC-GPTVHQCSTNTYKGLCEL 109
Db 75 SI-SPDIACO-PVEHICMPKRGYQGVTLVGNHNGVLVCIQVQARKFRSLNSELTGACSL 132
Qy 110 FGSMLRQOPKPPPALRG-----C-----PQE 131
Db 133 LTPNLDLQAAYFSDLEGFLDPGAHVDSGDYCRSGKSGTGEKKSARRRRTVBEDEED 192
Qy 132 DSDIATFLDGSGLIIPHDPRMELVSTIMEQL--KSKTFLSLMOYSEPRHPTKEF 189
Db 193 GTEIATVDDSGSGISGSDFOKAKFISTMMRNPFYKCFECNFALVQVGAVIQTEFLQES 252
Qy 190 QNNPNRSLIKPIITQLGRTHRTATGLKRVVRELFNITNGARKAKFKILFLITDGEKFGDP 249
Db 253 RDINASLAKVQSIQVQKVTIKASQHVLDNIFIPSGRSGKALKVMVLLTDGIDFGDP 312
Qy 250 LGYEDVTPELDREGVIRVYVLFQDAPRSEKSRQELNIVASKPRADHVFOANNFEALKTQV 309
Db 313 LNLTVINSQKQGVVFAIGVGRPKNNYRELKLIASDPKBAHTFKVTNYSALDGLL 372
Qy 310 NQLEKIFAEGTGTGSSSFEHMSQEGFSAITNGP--LLSTVGSYDWAGGVFLY-TS 367
Db 373 SKLQRIVEHGT---VGDALQYLAQTGFSAQLDKGVLLGTGVAFNWSGGALLYSTQ 429
Qy 368 KEKSTPINT-RVDS-DMDNAYLGYAAAILNRVQSLVIGAPRYQIHGVAMFRONTGM 425
Db 430 NGRCCFLNQTAKESRTVQSYLGSLAVLRKAGHSYVAGAPRHKLRGAVFLRKEDE- 488
Qy 426 WESNA--NNKGTQIGAYFGASLCSDVDNSNGSTDVLIGAPHYEYTRGQVSVCPPLR 482
Db 489 -EDAPYRRIEGHQWGSYFSGVLCVDIDMDCTDFTLLVAAPFHINGEGRVYVYQVPE 547
Qy 483 GQARWOCDAVLYGEGQCPGRFGAALTVLGDVNGDKLTQVAIGAP-----GEENRGA 536
Db 548 -QDASFSLAHTLSGHPGLTNSRFQFAMAAGVDINQDKFTDVAICAPLEGFGAGDGASYGS 606
Qy 537 VYLFHGTSGSGISPSHRSIAGSKLSPLOYFGSLGGQDLMGDLVLTGACGHVLL 596
Db 607 VYINGHSG-GLYDSPQQIETASSVASGLHYFGMSVGGLDFTNGDGLADITVSRDSAV 665
Qy 597 LRSQVILRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVQKS---TRDLREG 653
Db 666 LRSRPVVDLTVSMTFE-----DALPMVFIGK--DVNLCFEVDSVVVASEPGLREM 715
Qy 654 QIQSVWYTDALDGRSHRAVFNENKSTNRQTVLGLTQC-----696
Db 716 FLNFTVDV-----TKQRQLQCESSGCSCLRKWNGSGFICEHFWLI 760
Qy 697 ETLLQLPNCIEDPVSPIVLELNFSLVGTFLSAFNLN----PVLAEADAQRLTALF--P 750
Db 761 STEEL-----CEEDCFSNITIKVYE-----FQTSGGRRDYPNFTL--DHYKEPSAIFQLP 809
Qy 751 FEKNCNDNTQDDLSITFSFMSLDCLVVGCPREFNVTVTVRDNGDSYITQVTFPFLD 810
Db 810 YEKDCKNKVFCIAEIQUTN-ISOQLVGVTVKEVTNVIISLTNSGDSYNTNMAIYPRN 868
Qy 811 LSVKSVTLQNRQSRWRKLACASSTEVSGALKSCITSINEHPIPFENSEVFNTFDV 870
Db 869 IQFKKI-----QKPSVQCCDDPEV-----ASVLWNCKIGHPII-KRSSVNVSVTQWL 918
```

RESULT 8

ITAE\_HUMAN STANDARD; PRT; 1179 AA.

ID ITAE\_HUMAN AC F38570; Q9NZU9;

DT 01-OCT-1994 (Rel. 30, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1 antigen) (CD103 antigen) (Integrin alpha-IEI).

DE ITGAE.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]\_SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.

RC TISSUE=Leukemia, and Lymphocytes;

RX MEDLINE=94164962; PubMed=811947;

RA Shaw S.K., Cepek K.L., Murphy S.A., Russell G.J., Brenner M.B., Parker C.M.;

RA "Molecular cloning of the human mucosal lymphocyte integrin alpha E subunit. Unusual structure and restricted RNA distribution.";

RL J. Biol. Chem. 269:6016-6025(1994).

RN [2]\_REVIEWS TO 88-114.

RP Parker C.M.;

RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [3]\_SEQUENCE OF 53-1179 FROM N.A.

RP TISSUE=Retal kidney;

RX MEDLINE=20138496; PubMed=10673275;

RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G., Shotelersuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A., Green E.D.;

RA "The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion.";

RL Genome Res. 10:165-173(2000).

RN [4]\_MUTAGENESIS OF ASP-109 AND PHE-316.

RP MEDLINE=20400502; PubMed=10837471;

RX Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y., Brenner M.B.;

RA "The role of alpha and beta chains in ligand recognition by beta 7 integrins.";

RL J. Biol. Chem. 275:25652-25664(2000).

CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL CELL MONOLAYERS.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN  
CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED IN MUCOSAL  
CC EPITHELIAL CELLS.  
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -1- SIMILARITY: Contains 1 VWFA domain.  
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdi03.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L25851; AAB59359.2; -  
CC EMBL: AF168787; AAP43107.1; -  
CC PIR: A53213; A53213.  
CC HSP: P11215; IABX.  
CC Genew: HGNC:6147; ITGAE.  
CC MIM: 604682; -  
CC GO: GO:0008305; C:integrin complex; TAS.  
CC GO: GO:0004895; F:cell adhesion receptor activity; TAS.  
CC InterPro: IPR000413; Integrin\_alpha.  
CC Pfam: PF01839; FG-GAP; 3.  
CC Pfam: PF00357; integrin\_A; 1.  
CC Pfam: PF00092; vwa; 1.  
CC PRINTS: PR01185; INTEGRINA.  
CC PRINTS: PR00453; VWFA DOMAIN.  
CC SMART: SM00191; Int\_alpha; 3.  
CC SMART: SM0327; VWA; 1.  
CC PROSITE: PS00242; INTEGRIN\_ALPHA; 1.  
CC PROSITE: PS00234; VWFA; 1.  
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
CC Signal; Repeat; Polymorphism; Magnesium; Calcium.  
CC SIGNAL 1 18  
CC CHAIN 19 1179 INTEGRIN ALPHA-E.  
CC CHAIN 19 1177 INTEGRIN ALPHA-E LIGHT CHAIN.  
CC CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.  
CC DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 1125 1147 POTENTIAL.  
CC DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).  
CC DOMAIN 181 198 GLU-RICH (ACIDIC).  
CC REPEAT ? ? FG-GAP 1.  
CC REPEAT ? ? FG-GAP 2.  
CC DOMAIN 145 199 X-DOMAIN (EXTRA DOMAIN).  
CC DOMAIN 200 391 VWFA.  
CC REPEAT 401 456 FG-GAP 3.  
CC REPEAT 457 506 FG-GAP 4.  
CC REPEAT 510 571 FG-GAP 5.  
CC REPEAT 573 638 FG-GAP 6.  
CC REPEAT 641 693 FG-GAP 7.  
CC CA BIND 522 530 POTENTIAL.  
CC CA BIND 586 594 POTENTIAL.  
CC CA BIND 654 662 POTENTIAL.  
CC SITE 1150 1154 GREY MOTIF.  
CC DISULFID 70 79 BY SIMILARITY.  
CC DISULFID 126 159 BY SIMILARITY.  
CC DISULFID 706 762 BY SIMILARITY.  
CC DISULFID 823 829 BY SIMILARITY.  
CC DISULFID 893 907 BY SIMILARITY.  
CC DISULFID 1008 1033 BY SIMILARITY.  
CC DISULFID 1041 1057 BY SIMILARITY.  
CC CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 726 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 782 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 857 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 934 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 954 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 360 D -> E.  
FT VARIANT 1041 /FTID=VAR\_008884.  
FT VARIANT 1041 C -> S.  
FT MUTAGEN 208 D->A: LOSS OF B-CADHERIN BINDING.  
FT MUTAGEN 316 F->A: LOSS OF B-CADHERIN BINDING.  
FT CONFLICT 477 V -> I (IN REF. 3).  
FT CONFLICT 482 Q -> R (IN REF. 3).  
FT CONFLICT 950 R -> W (IN REF. 3).  
FT CONFLICT 1019 A -> V (IN REF. 3).  
SQ SEQUENCE 1179 AA; 13008 MW; E558902EDF9D95E1 CRC64;  
Query Match 19.2%; Score 1128; DB 1; Length 1179;  
Best Local Similarity 28.7%; Pred. No. 5.1e-68;  
Matches 338; Conservative 213; Mismatches 454; Indels 172; Gaps 39;  
QY 45 GSXYQDYSTGS--CEPI-RLQVP-----VAVNMSLGLSLAATSPOLLACGTVHQ 95  
DB 65 GFLHRCSLVQDEILCHPVEHPVPRGRHGVTVVRSHHGVLICI-----QVLVRRP--HS 117  
QY 96 TCSENYVYKGLCFLPGSNLRQPO-----SDAFIDGSGSIIPHDPRMKELVSTIMEQL-- 119  
DB 118 LSSELT---GTCSLLGPDLPQAOANFDEMLNLDPEARVDTGDCYSNKEGGEDDVNTA 174  
QY 120 KFEALRGCPQED-----SDFIADGSGSIIPHDPRMKELVSTIMEQL-- 164  
DB 175 RQRALKEEEDKEEEDDEEAEAGTEIATLDGSGSIDPDPDFORAKDFISNMVNFYE 234  
QY 165 KSKSTLFSLMQYSEEFRIHFTFKFQNNPNRSLKPTOLLGRTHHTATGLKVVRELEN 224  
DB 235 KCFECNFAIVQYGVQTEFLDRLSDQWASLARQNTQVGSVTKTASAMQVLDOSIT 294  
QY 225 ITNGARKAFKILFLITDGEKFGDPLGYEDVPELDREGVIRYVLGFGDAPFSEKREL 284  
DB 295 SSHGSRKASKVMVYLTDDGIFEDPLNLTTVINSPKMQGVRFALGVGFEFKSARTAREL 354  
QY 285 NTVASKPPDRHVQANFEALKTQVQLREKIPATECTGTGSSSSPEHEMSQEGESAAT 344  
DB 355 NLIASDPDDETHAFKVTNTYALDGLSLKLYNIISMEGT---VGDALHVLQAGIFSAQL 411  
QY 345 SNGP--LLSTVGSYDWAGGVFLY--TSKSKSTFINMTRVDSMDNDA---YLGVAAILRN 398  
DB 412 DERQVLLGAVGAFDWSGGALLYDTSRRGRFLNQTAADAADEAAQYSVLGVAVLHKT 471  
QY 399 RVQSLVLCAPRYQHI GLVAMFR--QNTGWESNANY--KGTQICAYFGASLCSVDVDSNGST 456  
DB 472 CSLSYVAGAPQYKHG--AVFELQKEGAEASPLVLEGEQMGSYFGSELCPVIDIDMGST 529  
QY 457 DLVILGAPHYTEQTRGGQVSVCPDRGORARQWQDAVLYGEOGQPMWRFGAALTJVLGVN 516  
DB 530 DFLVAAPFYVHVGEGRVYVTRLSE--QDGSFLARILSGHGFNAPFGFMAAGDL 588  
QY 517 GDLTADVAIGAP---GREDNR--GAVYLFHGTSGSGISPSHSQRIAGSKLSPLRYQFQ 570  
DB 589 QDKLTADVAIGAPLEGFGADOGASFGSVIYNG--HWDGLSASPSQRIASSTAVAPGLQYFGM 647  
QY 571 SLSCGGDLTMDGLVDLTVGAQGHVLLRLSQPLRVKAIWEFNPREVARNVFECNQVVK 630  
DB 648 SMAGFDISGGLADITVGTIGQAIVFRSRPVVRUKVSKAFTPSALP-----ICF 697  
QY 631 KEAGRVVRLVQKSTRDLREGQIQSVVVTVDLALDSGRPHSRVAVFNETKNSRTRQTVL 690  
DB 698 NGVNVRLCFEI--SSVTTASESGLEALLNFTLDVVGKRRRLQCCSDVRSCLGCLREWS 756  
QY 691 GLUTQCEVILKQLPN-----CTEDPVSPIVRLNFSVGTFLSAFGLNLRPLVLAEDQRLPT 746

```
Db 757 SCSQCELDL-LIMPTEGELCEEDCFNSASVKVSYQL-QTPGQTDHPQILDYRTEPPAI 814
Qy 747 ALPFFKNCNDNIQDDLSITFSFMSLDCLVVGGPRENVTVTVNDGEDSVRTQVTFP 806
Db 815 FQLPYKACQKFLFCVAELQLA-TTWSQCELVVGLTKELTLNLTNSGEDSVMTSMALN 873
Qy 807 FPLDLSVRKYSTIQMORSQSRWELACESASSTVSGALKSTSCSINHPIPPENSEVTENI 866
Db 874 YPRNLQ-----LKRQKPPSPNIQDDPPQV---ASVLIMNCRIGHPVL-KRSSAHVSV 923
Qy 867 TPVDVSKASLGNKLLKANYTSNN-----MPRTNKTEFQ---LPLPKVAVYVMTSHGV 919
Db 924 VMQLEENAPNRTADITVTVTNGNRRSLANETHTLQFRHGFVAVLSKPSIMYVNTGQGL 983
Qy 920 S--TKVLPNTASENTSRVMOHOVQVSNLQORSIPISIVPLVPRVIRMQTVWDRPQVTFSE 977
Db 984 SHKKEFLFHVHGEN---LFGAEYQ-----LQICVPTKLGQLQVAAVKLRTQ 1028
Qy 978 NLSTCHTKERLPSHSDFLAELRKAPVYVNCIAVCQRIQCDIPFFGIQEBFNATLKNLS 1037
Db 1029 ASVTCTWSQERACAYSS-VQVHEEHSVSCVIA-----SDKENVTVAABIS 1073
Qy 1038 PDWYIKTSHHLLIVST-----ABILFNDSVFTLLPGQGFVRSQTSKTVKVEFPVNP 1091
Db 1074 WD-----HSEELKDVTEQLIGETISFNKSLYEGLNENH--RTKITVTVFLKDKYHSL 1125
Qy 1092 PLVIGSVSGGLLALITAAALYKLGFPKROYKDMASE 1128
Db 1126 PIILKSGVGLLVILVILVLPKCGFKRYKQQLNLE 1162

RESULT 9
ID ITAL HUMAN STANDARD; PRT; 1151 AA.
AC P56139;
AT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=93155124; PubMed=8428973;
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
  1 subunit";
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-
CC E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN 1-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH 1-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAMB=PROW; NOTE=CD guide CD49a entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm".
DR PIR; A45226; A45226.
DR PDB; 1QCS; 17-MAY-00.
DR DR; GeneW; HGNC:6134; ITGA1.
DR DR; MIM; 192968; -.
DR DR; GO; 0008305; C: integrin complex; TAS.
DR DR; GO; 0004895; P: cell adhesion receptor activity; NAS.
DR DR; GO; 0005518; P: collagen binding; TAS.
DR DR; GO; 0007160; P: cell-matrix adhesion; NAS.
DR DR; InterPro; IPR000413; Integrin_alpha.
```

```
DR InterPro; IPR002035; VWFA_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium; 3D-structure.
FT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1114 1136 POTENTIAL.
FT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT 2 2 FG-GAP 2.
FT REPEAT 147 360 VWFA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GFFKR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 583F3C1AABF52808 CRC64;
```

```
Query Match 18.5%; Score 1085.5; DB 1; Length 1151;
Best Local Similarity 27.6%; Pred. No. 3.7e-65;
Matches 342; Conservative 209; Mismatches 488; Indels 201; Gaps 44;
Qy 1 FNLDTENAMTFCQENARG-FGQSVVOL---QGRVWVGAPQETVAANQRGSLYQCDYSTGS 56
Db 1 FNVDRVKNMTFGPVDEMGYTVQVYEBEGKWLIGSLGVQPKRNTGDVVKCPVGRGE 60
Qy 57 CEP-IRLQVPVEA-----VNSLGLSLAATSPPOLLAGCTPVHGTCTENTYVKGL 106
Db 61 SLPCVKLLDLPVNTSIPNVTVENKENTMFGSTL-VTNPNGGFLACGLYAYRCGLHYTTGI 119
Qy 107 CFLFGSNLEQQQPKFPEALRGCFQEDSDIAFLIDCGSGIIPHDFFRMKELVSTIMEQLKK 166
```

Db 120 CSDVSTPQVNSIAB--VQECSTQ-LDIVIVLDSGNSIYPWD-----SVTAFLNDLLKR 171  
Qy 167 -----SKTFLSMQYSEBRIIHTFKFQFNPNPRLIKPITQOLLGR-THETATGLKVR 220  
Db 172 MDIGPKQFQVGVQGENYTHFNLKYSSTSEVLVAACKIIVQRGQRTMFGTDTARK 231  
Qy 221 ELFNITNGARKNAKFLILITDGEFGDPLGYEDVPIELDRGVIRYVLGFGDAFR---- 276  
Db 232 EAFTEARGARRGVKKVMVITDGEH-DNHRLLKVIQDCEDENI QFSAIILGSYNGNL 290  
Qy 277 -SEKRSQELNIVASKPPDRHVQANNFEALKTQVQNLREKIPAIQSTQGTGSSSEHEHMS 335  
Db 291 STEKFVEEIKSASEPTEHFFNVSDALAVITVKTGLGERIFALEATADQSAASEMEMS 350  
Qy 336 QSGFSAIITNSGNPLLSITGVSVDWAGVFLYTSKE-----KSTF-INMTVDSDMDDAVILG 389  
Db 351 QTGFSNHYSDWMLGAVGAYDWGTVVMQKASQIIIPENTTFNVSTKKEPL-ASVLG 409  
Qy 390 YAAAILNRVQSL-VLGPAPYOHIGLVAMFQNTGMESNANVKGTQIGAYFGASLCSV 448  
Db 410 YTVNSATSSGDLVYIAGQPRYNTQVYIYRMEDGNIKILOTLGQEQIGSYFGSILTF 469  
Qy 449 DVDNSGTDVLIGAPHY-----YECYR-GGOVSVCPLEPGQARQWQDA 492  
Db 470 DIDKDSNTDILLVADPMYMGTEKEGKVYVVALMQTRFEYQMSLEPIKQTCSSRQHNS 529  
Qy 493 VLYGEQGPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLPFGTSGSGISPS 551  
Db 530 CTTENKNEPCGARFGTJAAVADLNDLNDGVNDVIGAPLEDHGGAVYVHG-SGKTIKRE 588  
Qy 552 HQRIAGSKLSPQLQVFGSLSGGQDLTDGLVDLTGVAQGVLLRSPQVLRKAIMEF 611  
Db 589 YAQRPISGGGKTLKFPQSGSIHGMELNGDGBDVTDTIGLGAALFWSRDAVAVKVTWVF 648  
Qy 612 NPREVARNVPCNDQVVKKEAG--EVRVCLHVO-KSTRDLRREGQIQSVVYDLDLDSG 668  
Db 649 ENKVMIOKNCH--MEGKETVCINATVCFEVLKSKEDTIVEADIQ-----YRVLDSL 701  
Qy 669 RHRSRAVNET-----KSTRQTOVLGTQTCETLKLQNLNCTEDPVPVILRLNLSL 722  
Db 702 RQISRSFSGTOERKQVQRTNITVKRSC-----TKHSFYMLDXHDFQDSVR---ITLDFNL 753  
Qy 723 VGTLSAFGNLAPVLAEDAQLFTALFPFEKNGNDNICODLSLITSPSMSLDCLVVGPP 782  
Db 754 T-DPENG-----PVLDDSLPNSVHEIYIPAKCGKKEKISLDSLVATEKDLIVRSQ 807  
Qy 783 RE-FNVTVTRNDGSDSYRTQVFFPLPLSLYRKSTLQNRQSRWSRLACASSTEVS 841  
Db 808 NDKFNVLTVKNTKDSAYNTRTIVHYSPLVSPGTEAIQKD-----SCBSN----- 853  
Qy 842 GALKSTCSINHIPIFENSEVTFNITFDVDSKASLGN-KLLAKANVTSENMPRTNKTEF 900  
Db 854 ---HNITCKVGPFLRRGBWVTKILFQNTSYLMENVTIYLSATSDSEPPETUSDNVV 910  
Qy 901 QLELPVKYAVVWVTVSHGVSTKYNFTASENTSRVMHQYQVSN-----LGQRS----- 949  
Db 911 NTSIPVKYEVGLQFYS-SASEYHISAETNTPVEVINSTEDIGNINIFYLIRKSGSPM 969  
Qy 950 --LPLSLVF-----LVPRLNQTIVDRPQVTFSENLSSTCHYE-----R 988  
Db 970 PELKLSISFPNMTSGPYLYPTGLSS-----SENANCRPHIFEDPFSINSKGL 1018  
Qy 989 LPSHSDFLAELAKAPVNVCSIAVCQRIQCDIPFGIOE-----EFNA 1030  
Db 1019 MTTSTD---HLRGITLDCNTKCFATITCNLTSSDISQVNVSLILWKPTFKSYFSSLNL 1075  
Qy 1031 TLKGNLSDFWYIKTSRNNHLLIVSTABILFNDVSFTLLPQGGAFVRSQVETKVEPEVNP 1090  
Db 1076 TIRGEL-----RSENASLVLSSN-----OKRELAIQISKDLGPR 1111  
Qy 1091 LPL--TVGSSVGLLLALITLALYKLGFFKCYKDMSE 1128

Db 1112 VPLWVILLSAPAGULLMLLILALWKIGFFKRPPLKXKMEK 1151

RESULT 10

ITAH HUMAN STANDARD; PRT; 1189 AA.

AC Q9UKX5; Q9UKQ1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-11 precursor.

GN ITGAI1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

FN [1];

SEQUENCE FROM N.A.

RP TISSUE=Fetal heart, and Osteoblast;

RC TISSUE=Fetal heart; PubMed=10486209;

RX MEDLINE=994117678; PubMed=10464311;

RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;

RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D., Wang S.-X., Morris C.M., Krissansen G.W.;

RA "Cloning, sequence analysis, and chromosomal localization of the novel human integrin alpha11 subunit (ITGAI1).";

RT Genomics 60:179-187(1999).

RL [2]

SEQUENCE FROM N.A.

RP TISSUE=Petal muscle, and Uterus;

RC TISSUE=Petal muscle; PubMed=10464311;

RX MEDLINE=99395147; PubMed=10464311;

RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;

RA "cDNA Cloning and Chromosomal Localization of Human alpha(11) Integrin. A collagen-binding, i domain-containing, beta(1)-associated integrin alpha-chain present in muscle tissues.";

RT J. Biol. Chem. 274:25735-25742(1999).

RL [3]

SEQUENCE OF 954-1188 FROM N.A.

RP TISSUE=Fibroblast;

RC Andreu N., Estivill X., Escarceller M., Sumoy L.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11 ASSOCIATES WITH BETA-1.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS, SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.

CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING FETAL MUSCLE CELLS (IN VITRO).

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 1 VMFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: AF109681; AAF01258.1; -

CC EMBL: AF137378; AAD51919.2; -

CC EMBL: AL359064; CAB94392.1; -

CC HSPG: P17301; IAOX.

CC Genew: 604789; -

CC MIM: 604789; -

CC GO: GO:0008305; C:integrin complex; TAS.

CC GO: GO:0004895; F:cell adhesion receptor activity; TAS.

DR	GO: GO:0005518; F:collagen binding; TAS.	QY	1	FNLDTENAMTFOENARG--FGOSVVO--LOGSR-VVVGAPQETVAANQSGSLYCCDYSTGS	56
DR	GO: GO:0007150; P:cell-matrix adhesion; TAS.	DB	23	FNWDTKRPVTPGSRATFAFGYTVQOHDISGNKWLWVGAPLETNGYKTDGVDYKCPVIHGN	82
DR	GO: GO:0007517; P:muscle development; TAS.	QY	57	CEPTRL-----QVPVEAVNMSLGLSLAATTSPFOLLACGPTVHOTCSNTTAVKGLCFLF	110
DR	InterPro; IPR000413; Integrin_alpha.	DB	83	CTKLNLRVTLNVSERKONMRLGLSLATNPKNDSFLACSPWSHSECGSSYVTTGMSRV	142
DR	InterPro; IPR01839; FG-GAP; 3.	QY	111	GSNLRQOPKPPALRGCPEDSDIAFLIDSGSIIPHDFRMKELVSTIMQLKK----	166
DR	Pfam; PF00092; vwa; 1.	DB	143	NSNFRPSKTAP-ALQRC-QTYMDIVI VLDGNSIYP-----WVEVQHFLNLIKRPYIG	195
DR	PRINTS; PR01185; INTEGRINA.	QY	167	-SKTLFSLMQYSEEFRIIHTFKEFQNNPRSLIKPITOLLG-RTHTATGLRKVVRELFN	224
DR	SMART; SM00191; Int_alpha; 5.	DB	196	PGQIQGVVQYGEDVVEHFLNDRSVKDVVEAASHIEQGGTETRTAFGISPARSEAFQ	255
DR	SMART; SM00327; vwa; 1.	QY	225	ITNGARKNAFKILFLITDGEKPEDPLGYEDVITPELDREGVIRY----VLGPGD--AFRSK	279
DR	PROSITE; PS00242; INTEGRIN ALPHA; FALSE_NEG.	DB	256	--KGRKGAKKYMITVITDGHSDSP-DLEKVIQOSERDNTVYAVAVLGYNNRGINPET	312
DR	PROSITE; PS0234; VWFA; 1.	QY	280	SQELMTVASKPPROHVFOANNFEALKTVQNLREKIPIAIEGTOTGSSSSPHEHMSQEGF	339
DR	InterPro; IPR000413; Integrin_alpha.	DB	313	FLNEIKYIASDDDKHFFNVTDAAALKDVIDALGRIPLSECTNK-NETSFGLWSQTGF	371
DR	PRINTS; PR00453; VWFADOMAIN.	QY	340	SAATSNGLSTVSGSYDWAGGVFLYTSKEK-----STPNNTRVDSMDNDAYLGHAAAI	394
DR	SMART; SM00191; Int_alpha; 5.	DB	372	SSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKNHKGAYLGYTVTS	431
DR	PRINTS; PR00453; VWFADOMAIN.	QY	395	IIRNEV-QSLVLAGAPRYOHIGLVAMP-RQNTGMSNSANVKGTQICAYFGASLCSVDVS	452
DR	SMART; SM00191; Int_alpha; 5.	DB	432	VVSSQGRVYVAGAPRFNHTGKVIILFTHNNESLIHQMRGQQIGSFSEITSVDIIG	491
DR	PROSITE; PS00242; INTEGRIN ALPHA; FALSE_NEG.	QY	453	NGSTDVLIGAPHYYEQTR-GQVSVCPPLPRQARWQCDVLYGEGQGPWGRFGAALTV	511
DR	PROSITE; PS0234; VWFA; 1.	DB	492	DGVTVDLLVGAPMYENEGRGKVVYEL---RQNFVYVNGTLKDSHYSQYARFGSSIAS	548
DR	InterPro; IPR000413; Integrin_alpha.	QY	512	LGDVNGDKLTVAIGAPGEEDNRGAVLPHGTSGSGISPSHSGIAGSKLSPLOYFGOS	571
DR	PRINTS; PR00453; VWFADOMAIN.	DB	549	VRDLNQDSYNDVVVGAPLEDNHAGAIYIPHGPRGS-ILKTPQRTASLALGLQIPGCS	607
DR	SMART; SM00191; Int_alpha; 5.	QY	572	LSGGQDLTMDGLVDLTVGAGQHVLRLSQPLRVKAIKMFNPREVARNVF--ECNDQVVK	629
DR	PROSITE; PS00242; INTEGRIN ALPHA; FALSE_NEG.	DB	608	IHQDLNEDGLIDLAVGALGNVILWSRPVQINASHLHFPSKI--NIFHRDC-----	659
DR	PRINTS; PR00453; VWFADOMAIN.	QY	630	GXEAGEVRVCL-----HVQKSTRDLREBQIQSVVYTDALDSCGPHSRVAF	676
DR	SMART; SM00191; Int_alpha; 5.	DB	660	-KRSGRDATCLAAFLCFTPIFLAPHFTTVG-----IRYNATMDERRYTPRAHL	708
DR	PROSITE; PS00242; INTEGRIN ALPHA; FALSE_NEG.	QY	677	NETKNS-TRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLELNFSLVGTPLSAFGLNRP	735
DR	PROSITE; PS0234; VWFA; 1.	DB	709	DEGGDFTRAVLLSSGQELCERINFHVL-DADYVKVTPSVSEYSLDP-----DHGP	761
DR	InterPro; IPR000413; Integrin_alpha.	QY	736	VLADARLFTALPPEKNCNDNTCQDDL-----SIT	768
DR	PRINTS; PR00453; VWFADOMAIN.	DB	762	MLDDGHPPTLRVSVPPWNGCNEDEHCVPLVLDARSDDLPTAMEYQCVLRKPAQCSAYT	821
DR	SMART; SM00191; Int_alpha; 5.	QY	769	PSFMSLDLWVGSPREFNVTVTRNDGDSYTVTFFFPPLDLSYRKVSTLQNSQSRW	828
DR	PROSITE; PS00242; INTEGRIN ALPHA; FALSE_NEG.	DB	822	LSFDITVPIESTRQVAVATELENGENAYSTVLNISQSANLQF--ASLIQKEDSDGS-	878
DR	PRINTS; PR00453; VWFADOMAIN.	QY	829	RLACSASTVSGALKSTSCSINHPIIPENSEVTFNTITFVUDSKASLGNKLLKANVTS	888
DR	SMART; SM00191; Int_alpha; 5.	DB	879	-IECVNEER-----RLQKQVCNVSYPFPAKAKVAFRLDFEF-SKSIHLHLEIELAAS	931
DR	PROSITE; PS00242; INTEGRIN ALPHA; FALSE_NEG.	QY	889	ENNMPTNKTE--FOLELPVKYAVVMVVTSHGVSTKY---LNPTAS--ENTSRVMQHQVQ	941
DR	PROSITE; PS0234; VWFA; 1.	DB	932	DSNERDSTEDNVAPLRFILKYADVLFTRSSLSHYEVLKNSLERYDGIQFPFSCIFR	991
DR	InterPro; IPR000413; Integrin_alpha.	QY	942	VSNLQO---RSLPISLVFLVPLVRLNQTVIWRDQVTFSENLSSTC-----HTKERLP	994
DR	PRINTS; PR00453; VWFADOMAIN.	DB	992	IQMLGLFPIHGMKTIPIATRSNGNLLKLKDLFT-DEVANTSCINMGNSYRTPVPE	1050
DR	SMART; SM00191; Int_alpha; 5.	QY	995	FLAELEKAPVNCISIAVCQRIQCDIPFFGIQEEFNATKGNLSFDWY-----IKTSHNL	1049

Query Match 18.4%; Score 1079.5; DB 1; Length 1189;  
Best Local Similarity 28.3%; Pred. No. 9.8e-65;  
Matches 349; Conservative 212; Mismatches 501; Indels 173; Gaps 48;

Db 1051 --EDLRAPQLNHSNDVVSINCNILVP--NOEINPHILGNL---WLSIKALKYKSMKI 1104  
 QY 1050 LIVSTRAILLFNDVVFILLPQOGAFVRSQTETKVEPEVFN-----PPLIVGSSVGGILL 1103  
 Db 1105 MVNAALQROPH-SPF-----IFREDSRQIVFEISKQEDMQVPIIIVGSLTGLL 1155  
 QY 1104 LIALITAAALYKLGFFK--ROYKDMSEGGPGGAEP 1136  
 Db 1156 LLALLVLALWKLGGFFSARRRR-----PGLDLP 1183

RESULT 11  
 ITA2\_BOVIN STANDARD; PRT; 1170 AA.  
 AC P53710;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)  
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).  
 GN ITGA2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metazoa; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94193647; PubMed=7511592;  
 RA Kamata T., Puzon W., Takada Y.;  
 RT Identification of putative ligand binding sites within I domain of  
 RT integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).  
 RL J. Biol. Chem. 269:9659-9663(1994).  
 CC -I- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT  
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-B-R IN  
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE  
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX.  
 CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
 CC ASSOCIATES WITH BETA-1.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -I- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -I- SIMILARITY: Contains 1 VWFA domain.  
 CC -I- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; L25886; BAB59255.1; --  
 DR PIR; I45914; I45914.  
 DR HSP; P17301; IAOX.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF A.  
 DR Pfam; PF01839; FG-GAP; 3-  
 DR Pfam; PF00357; integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00191; int\_alpha; 5.  
 DR SMART; SM00327; vwa; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.  
 NON\_TER

FT SIGNAL <1 18  
 FT CHAIN 19 1170  
 FT DOMAIN 1121  
 FT TRANSMEM 1122  
 FT DOMAIN 1144  
 FT REPEAT 34 92  
 FT REPEAT ?  
 FT DOMAIN 177 367  
 FT REPEAT ?  
 FT REPEAT 423 475  
 FT REPEAT 477 538  
 FT REPEAT 540 599  
 FT REPEAT 604 656  
 FT CA\_BIND 488 496  
 FT CA\_BIND 552 560  
 FT CA\_BIND 616 624  
 FT SITE 472 474  
 FT SITE 1146 1150  
 FT DISULFID 72 81  
 FT DISULFID 669 726  
 FT DISULFID 778 784  
 FT DISULFID 854 865  
 FT DISULFID 1008 1039  
 FT DISULFID 1044 1049  
 FT CARBOHYD 94 94  
 FT CARBOHYD 101 101  
 FT CARBOHYD 332 332  
 FT CARBOHYD 421 421  
 FT CARBOHYD 449 449  
 FT CARBOHYD 464 464  
 FT CARBOHYD 688 688  
 FT CARBOHYD 748 748  
 FT CARBOHYD 945 945  
 FT CARBOHYD 1063 1063  
 FT CARBOHYD 1070 1070  
 FT VARIANT 580 580  
 FT VARIANT 588 588  
 FT VARIANT 725 725  
 SQ SEQUENCE 1170 AA; 128929 MW; BECEF1C5F2448FB1 CRC64;

Query Match 18.2%; Score 1071; DB 1; Length 1170;  
 Best Local Similarity 27.5%; Pred. NO. 3.6e-64;  
 Matches 334; Conservative 218; Mismatches 495; Indels 168; Gaps 48;  
 QY 1 ENLDTENAWTQ-ENARFGQSVVQL---QGRVVVGAPOEIVAAHQSGSLVQC--DYST 54  
 Db 19 NVGLPKAKIFSGPSSQFGYAVQQFIPNKGWLLVGSFGFPKRMGVDYKCPVDIST 78  
 QY 55 GSCEPIRLQ-----VFVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTVYKGLC 107  
 Db 79 TTCEKLNLOTSTMSNVTEMKTNMSLGLTLRNVTGGFLTCGLPWAQCGSQYVTTGVC 138  
 QY 108 FLPGSNLQOQKPEALRGCPQEDSDIAFLDQSGSIIPHDPRMKELVSTIMEQLK-- 165  
 Db 139 SDVSPDF-QLRTSFAPAQTCP-SFIDVVVCDSENSIYPWD--AVKMFLEKFGVGLDIG 194  
 QY 166 KSKTLFSLMQYSEEPRIHFTFKBFONNPNRSLIKPIITQLL----GRTHATGLKVVRE 221  
 Db 195 PTKOMGLIQANNPRVFNLTNFKSKD--EMIKATSTQTYGGDLNTPKAIQYARDT 251  
 QY 222 LFNITNGAKNAFKILFLITDGEKFGDPGLGVEDVPELDREGVIRY---VLGF--GDAFR 276  
 Db 252 AYSTAAGGRPGATKVMVVVYTDGESH-DGSKLKAVIDQCCKNHLRFGVIAVGLYLNALD 310  
 QY 277 SEKSQELNTVASKPPRDHVFQANNFEALKTQVQNLREKI PAIEGTQTGSSSSPHEMSQ 336  
 Db 311 TKNLKEIKATASIPTEHFPNVSDPADLLEKAGTIGQIFSEGTVOG-GNFMQMSQ 369  
 QY 337 EGFSAAIT--SNGPLLSVGVSDMAGGVFLTSKSKTSPIFNT--RVSDMN-DAYLGA 391  
 Db 370 VGFSAEYSPQNILMLGAVGAYDMSGTVQKTPHGLHIFSKQAFRQILQDRNHSSYLGYS 429  
 QY 392 AAILLRNVQSLVLCAPRYQHIGLVAMFRQNTGMESNANV-----KGTQIGAYFGASL 445

```
Db 430 VASISTGNSVHFAAGPANTTGGIVLISN-----ENGWTVTQSOGQIGSYFGSVL 484
Qy 446 CSVDVDSKSTDLVLICAPHYEQR--GGQVSCVPLPRGORARQCDAVLYGQGPWG 503
Db 485 CAVDVNDOTIDVLLAGAPMTMDLKKKEGRVYLFTITKG-ILNWH--QFLEGPNGLENA 541
Qy 504 RFGAALTVLGDNVGDGLTVAIGAPGREDNAGAVLFGTSGSISPSHSORLAGS--KL 561
Db 542 RFGSAITALSIDINMGDNVIVGSPLENONGAVIYINGHEGM-IRLRYQKILGSDRAP 600
Qy 562 SPRLQYFQSSLSGGQDLTMDGLVDLTGAGQGHVLLRSQPVLRVKAIMEFNPREVARNVF 621
Db 601 SSHLYQFGRSLDGYGLDNGSDITDVSVCAPGVQVQLWSQSIADSVDSPTPKKI--TL 658
Qy 622 ECNDQVVKKAGAEVRCVLAVQKSTRDLRGQIQSVVTVYDLAD----SGRPHRAVEN 677
Db 659 KNAEI-----KKLCF-----SAKFRFTQNNQVRAIVNITIDBQFSRVSRLGFK 707
Qy 678 ETKNSTRRQTVGLGTQCE--TLKLQPNCEIDPVPSPVILRLNFSL--VGTPLSAPGNL 733
Db 708 ENNERCLQKTIWVSQACRCSEVIIHQPS-----DIISPLNLCWMSLENPGT----- 756
Qy 734 RPLVLAEDAQRLFTALPPEKXKCNNDICQDLSITP-----SPMSLDCLVGCPREFNVTV 789
Db 757 NPALVAYSETVKVFSIPFHQCGDGVCSIDLVLNVOQLPATQOQPFIVSNQKRLTFV 816
Qy 790 TVRNDGDSVYQVTFPPFLDLSYRKVSTLQNRQSQRSLACSSAST-EVSGALKSTS 848
Db 817 QLKXKESAYNTEIVDSENLFF-----ASMSMPVDGTEVTCQIASQKSVT 864
Qy 849 CSNIHPPIENSEVNTFPDVSQASLGKLLKANVTSENMPRTNKTFFQLSLPKVY 908
Db 865 CNVGPALKSKQVTFITFDNLQ-NLQNASISPRALSBSQENMADNSVNLKSLLY 923
Qy 909 AVYVVTSHGVTYKVLNFTASNTSRVQHOVQVSNLQZ-----SLPISLVLV 958
Db 924 DAEIHT--ASTNINPVEVSLDGNSSV--HSEF--DIGPKPIFKVITGVSVPVSM-- 976
Qy 959 PVRLNQTVIMDRPQVTFSEN---LSSCTHTKE-----RLPSHSDFTLAE- 998
Db 977 -----SVIIHPQVTKDKNPLVLTGVHTDQAGDISCEABINPLKIQOTSSVSFKSEN 1030
Qy 999 LKAPVNCSTAVCORICDIPFGIQEFNATLKGSLFDFWYIKTSHHLLIVSTAEI- 1057
Db 1031 FRHIELNCRATSCSNICWLRDLQKGEYFLNVSTRIWNGTFPAASTQVQLTAALID 1090
Qy 1058 LFNDSVFTL-----LPQGGAFVRSQTETKVEPFE-VENPLPLIVGSSVGLLLALITA 1110
Db 1091 TYNPQIYVIBENTVIP-----LTINKPEKVEVPTGVIVGSVIAGILLALVA 1140
Qy 1111 ALYKLGPFKRYKDM 1125
Db 1141 ILWKLGFPRKRYKM 1155

RESULT 12
ITR2 MOUSE
ID ITR2 MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;
```

```
RX MEDLINE=94363406; PubMed=8081889;
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
RA Banjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.,
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RL not virus binding.";
RL Cell Adhes. Commun. 2:131-143(1994).
RN [2]
RP SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.B., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RL 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -I- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -I- SIMILARITY: Belongs to the integrin alpha chain family.
CC -I- SIMILARITY: Contains 1 VWFA domain.
CC -I- SIMILARITY: Contains 7 FG-GAP repeats.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z29987; CAA82877.1; --
DR EMBL; X75427; CAA53178.1; --
DR PIR; S44142; S44142.
DR HSP; P17301; IAOX.
DR MGD; MGI:96600; Itga2.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP_3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00327; VWA; 1.
DR SMART; SM00191; Int_alpha; 4.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
DR Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1
FT CHAIN 26
FT DOMAIN 27 1178 INTEGRIN ALPHA-2.
FT TRANSMEM 1130 1151 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1152 1178 POTENTIAL.
FT REPEAT 42 100 CYTOPLASMIC (POTENTIAL).
FT REPEAT ? 2 FG-GAP 1.
FT REPEAT 185 375 FG-GAP 2.
FT REPEAT ? 3 FG-GAP 3.
FT REPEAT ? 4 FG-GAP 4.
FT REPEAT 431 483 FG-GAP 5.
FT REPEAT 485 546 FG-GAP 6.
FT REPEAT 548 607 FG-GAP 7.
FT REPEAT 612 664 POTENTIAL.
FT CA_BIND 496 504 POTENTIAL.
FT CA_BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 GPFR MOTIF.
FT DISULFID 80 89 BY SIMILARITY.
```

FT	DISULFID	677	734	BY SIMILARITY.	
FT	DISULFID	786	792	BY SIMILARITY.	
FT	DISULFID	862	873	BY SIMILARITY.	
FT	DISULFID	1016	1047	BY SIMILARITY.	
FT	DISULFID	1052	1057	BY SIMILARITY.	
FT	CARBOHYD	102	102	N-LINKED (GLNAC. . .)	(POTENTIAL)
FT	CARBOHYD	109	109	N-LINKED (GLNAC. . .)	(POTENTIAL)
FT	CARBOHYD	429	429	N-LINKED (GLNAC. . .)	(POTENTIAL)
FT	CARBOHYD	457	457	N-LINKED (GLNAC. . .)	(POTENTIAL)
FT	CARBOHYD	472	472	N-LINKED (GLNAC. . .)	(POTENTIAL)
FT	CARBOHYD	696	696	N-LINKED (GLNAC. . .)	(POTENTIAL)
FT	CARBOHYD	1054	1054	N-LINKED (GLNAC. . .)	(POTENTIAL)
FT	CARBOHYD	1071	1071	N-LINKED (GLNAC. . .)	(POTENTIAL)
FT	CARBOHYD	1078	1078	N-LINKED (GLNAC. . .)	(POTENTIAL)
SQ	SEQUENCE	1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;			
Query Match					18.0%; Score 1057; DB 1; Length 1178;
Best Local Similarity					27.8%; Pred. No. 3.2e-63;
Matches 341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;					
QY	1 FNLDTENAMTFQ--ENARFGQSVVQL---QGSRRVVGAPQEIIVAAQNRGLYOC---DYST 54				
DB	27 YRVGLPGAKIFSGPSSQFGYSVQQLTNPQGNLLNGSPMSFPENRMDGVYKCPVDLPT 86				
QY	55 GSCEPIRLQ-----VPVEAVNMSGLSLAAATSPPLACAGPTVHQTCSENTYVKGLC 107				
DB	87 ATCEKMLQNSASISNTEIKTNMSLGLTITRAPGTCGFLTCGLPWAHQCGNQYATGIC 146				
QY	108 FLFGSNLRQOPQ---KPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKELYSTIMBQL 164				
DB	147 ----SDVSPDFQFLTSFSPAQAQCSL-VDVVVVCDSENSIYP--MEAVKNFLVKFVTGL 199				
QY	165 K--KSKTFLSILMOYSEBFIHFTFKFQNNENPSRLIKPITQLLG-RTHATATGLRKVVRE 221				
DB	200 DIGPKTKOVALIOYANEPRIIFNLNDFETKEDVMQATSETROKGGDLTNTFRAIEFARDY 259				
QY	222 LFNITNGARKNAFKILLLTIDGERFGDPLGYEDVIPELDREGVIRY---VLGF--GDAER 276				
DB	260 AYSQTSQGRPGCATKVMVVMVVDGESH-DGSKLTKTVIQCNDDIELRFGIAGVLGVLNRLALD 318				
QY	277 SEKSRQELNVTASPRDRHVFQANFALKTVQNLBEKIFAEGTGTGSSSPHEHMSQ 336				
DB	319 TKNLIKRIKAIATPTERYFPFNVADEALKEKAGTLGEQIFSGTGTGQV--GDMFQMEMAQ 377				
QY	337 EGPSA---AITSNGPLLSYDNGAGGVELYTSKSTFTINMT--RYVDSMN-DAYLGIA 391				
DB	378 VGFSAADYAPQNDILMLGAVGAFDMSGTLVQETSHKPVIFPKQAFDQVLQDNHSSFLGYS 437				
QY	392 AATILNRVQSLVLGAPRYQHIGLVAMPFRONTGWNESNANY----KGTQICAGFASGLCS 447				
DB	438 VAAISTEDGVHPVAGAPEANTGTQIVLYSVNK---QGNVTYIQSHRGDQIGSYFGSVLCS 494				
QY	448 VDVDSNGSTDVLVLCAPHYBQTR--GGQVSVCPPLPGORARWQCDAVLYGEOQPWGRF 505				
DB	495 VDVKDITDVLVLCAPHYBQTR--GGQVSVCPPLPGORARWQCDAVLYGEOQPWGRF 551				
QY	506 GAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSQSRIAGSKLSPR- 564				
DB	552 GSAIAALSDINMDGFNDVIGSPVENESGAVIYNGHQGT-IRTKYSQKILGNSGAFRR 610				
QY	565 -LOVFGQSLSCGQDLTDGLVLTGAGHVLRLRSOPVLARVKALMEFPNPREVARNVPEC 623				
DB	611 HLQFFRSLGSLGDLNGSDITDVSIGALGVQIQLWSQSIAQVADVALEAFTP----- 660				
QY	624 NDQVVKGEAGEVRVCLHVQKSTDRLRREGQIQSVVTVYDLDALD----SGRPHSRAVFNET 679				
DB	661 -DKITLLINKAKITLKLCFRAEPFAGQNNQV--AILFNMTLDADGHSSRVTSRSGVREN 717				
QY	680 KNSTRQTVQLGLTQCTET--LKLQLPNCIEDPSPVIRLNFSLVGTPTLSAFGNLBPVL 737				
DB	718 SERFLQKNVAVYVQKCEHHSISIKPS---DVVNPLDLRVDISLENPGTS-----PAL 768				
QY	738 ABDAQRLFTALFPPEKNGNDNIQDLDLSI-----TPSFMSLDCLVVGGPREFNVNT 788				
RESULT 13					
ITA2	HUMAN				
ID	ITA2_HUMAN	STANDARD;	PRT;	1181 AA.	
AC	PI7301;				
DT	01-AUG-1990 (Rel. 15, Created)				
DT	01-AUG-1990 (Rel. 15, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)				
DE	(Collagen receptor) (VLA-2 alpha chain) (CD49b).				
GN	ITGA2.				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.				
RC	TISSUE=Endothelial cells;				
RX	MEDLINE=89308879; PubMed=2545729;				
RA	Takada Y., Hemler M.E.;				
RT	"The primary structure of the VLA-2/collagen receptor alpha 2 subunit				
RT	(platelet GP1a): homology to other integrins and the presence of a				
RL	possible collagen-binding domain.";				
RL	J. Cell Biol. 109:397-407(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Rieder M.J., Arwel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,				
RA	Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;				
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.				
RX	MEDLINE=98019223; PubMed=9353312;				
RA	Emsley J., King S.H., Bergelson J.M., Liddington R.C.;				
RT	"Crystal structure of the I domain from integrin alpha2beta1.";				
RL	J. Biol. Chem. 272:28512-28517(1997).				
RN	[4]				
RP	VARIANT HPA-5 (BR).				
RX	MEDLINE=94043762; PubMed=79012136;				
RA	Santos S., Kab R., Wanka M., Kiefel V., Mueller-Eckhardt C.,				
RA	Newman P.J.;				
RT	"The human platelet alloantigens Br(a) and Br(b) are associated with a				
RT	single amino acid polymorphism on glycoprotein Ia (integrin subunit				

alpha 2).";  
J. Clin. Invest. 92:2427-2432(1993).  
[5]  
VARIANT GIU-534.  
MEDLINE=20206009; PubMed=10744142;  
Kroll H., Gardemann A., Rechter A., Haberbosch W., Santoso S.;  
"the impact of the glycoprotein Ia collagen receptor subunit A1648G  
gene polymorphism on coronary artery disease and acute myocardial  
infarction.";  
Thromb. Haemost. 83:392-396(2000).  
-1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND B-CADHERIN. IT  
RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-R-R IN  
COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE  
EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
EXTRACELLULAR MATRIX.  
-1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2  
associates with beta-1. Interacts with HPS5.  
-1- SUBCELLULAR LOCATION: Type I membrane protein.  
-1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
-1- POLYMORPHISM: Position 534 is associated with platelet-specific  
allotypism HPA-5 (BR). HPA-5A/BR(A) has lys-534 and HPA-5B/BR(B)  
has glu-534. HPA-5B is involved in neonatal alloimmune  
thrombocytopenia (NAIT or NATP). The K534E polymorphism may play a  
role in coronary artery disease (CAD).  
-1- SIMILARITY: Belongs to the integrin alpha chain family.  
-1- SIMILARITY: Contains 1 VWFA domain.  
-1- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;  
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49b.htm".  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X17033; CAA34894.1; --  
EMBL; AF512556; AAM34795.1; --  
PIR; A33998; A33998.  
PDB; LAOX; 25-NOV-98.  
PDB; 1DZ1; 02-AUG-01.  
Genew; HGNC:6137; ITGA2.  
MIM; 192974; --  
GO; GO:0008305; C:integrin complex; TAS.  
GO; GO:0005886; C:plasma membrane; TAS.  
GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
GO; GO:0005518; F:collagen binding; TAS.  
GO; GO:0007596; P:blood coagulation; TAS.  
GO; GO:0007160; P:cell-matrix adhesion; TAS.  
GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
InterPro; IPR000413; Integrin\_alpha.  
InterPro; IPR002035; VWFA.  
Pfam; PF01839; FG-GAP; 3.  
Pfam; PF00357; Integrin\_A; 1.  
Pfam; PF00092; vwa; 1.  
SMART; SM00191; Int\_alpha; 5.  
SMART; SM00327; VWFA\_1  
PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
PROSITE; PS50234; VWFA; 1.  
Intergrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;  
3D-structure. 1 29  
CHAIN 30 1181 INTEGRIN ALPHA-2  
DOMAIN 30 1132 EXTRACELLULAR (POTENTIAL).  
TRANSMEM 1133 1154 POTENTIAL.  
DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).  
DOMAIN 1155 1161 INTERACTION WITH HPS5.

REPEAT 45 103  
FT REPEAT 188 ?  
FT DOMAIN 378  
FT REPEAT 433  
FT REPEAT 434  
FT REPEAT 488  
FT REPEAT 551  
FT REPEAT 610  
FT REPEAT 635  
FT REPEAT 667  
FT CA\_BIND 507  
FT CA\_BIND 563  
FT CA\_BIND 627  
FT SITE 1157  
FT DISULFID 83  
FT DISULFID 680  
FT DISULFID 795  
FT DISULFID 865  
FT DISULFID 1019  
FT DISULFID 1055  
FT CARBOHYD 105  
FT CARBOHYD 112  
FT CARBOHYD 343  
FT CARBOHYD 432  
FT CARBOHYD 460  
FT CARBOHYD 475  
FT CARBOHYD 699  
FT CARBOHYD 1057  
FT CARBOHYD 1074  
FT CARBOHYD 1081  
FT VARIANT 534  
FG-GAP 1.  
FG-GAP 2.  
VWFA.  
FG-GAP 3.  
FG-GAP 4.  
FG-GAP 5.  
FG-GAP 6.  
FG-GAP 7.  
POTENTIAL.  
POTENTIAL.  
GFFER MOTIF.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
N-LINKED (GLCNAC. .) (POTENTIAL).  
N-LINKED (GLCNAC. .) (POTENTIAL).  
N-LINKED (GLCNAC. .) (POTENTIAL).  
N-LINKED (GLCNAC. .) (POTENTIAL).  
N-LINKED (GLCNAC. .) (POTENTIAL).  
N-LINKED (GLCNAC. .) (POTENTIAL).  
N-LINKED (GLCNAC. .) (POTENTIAL).  
N-LINKED (GLCNAC. .) (POTENTIAL).  
N-LINKED (GLCNAC. .) (POTENTIAL).  
K -> E (IN ALLOANTIGEN HPA-5B;  
dbSNP:1801106)  
/FTID=VAR\_003977.

TURN 170  
STRAND 173  
TURN 183  
HELEX 188  
TURN 201  
STRAND 204  
TURN 206  
STRAND 209  
STRAND 220  
TURN 226  
HELEX 232  
TURN 241  
HELEX 252  
TURN 263  
HELEX 266  
TURN 269  
STRAND 275  
HELEX 289  
TURN 301  
TURN 302  
STRAND 304  
HELEX 313  
TURN 318  
HELEX 323  
TURN 331  
HELEX 337  
STRAND 341  
HELEX 353  
HELEX 354  
TURN 363  
SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match 17.9%; Score 1054; DB 1; Length 1181;  
Best Local Similarity 26.8%; Pred. No. 5.2e-63;  
Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;

QY 1 FNLDYENAMTFQ-ENARGFGQSVVQL---QGSRRVVVGAPQBIIVAAQNRSLYOC--DYST 54  
DB 30 YNVGLPEAKI FSGPSSEQFGYAVQQFINPKGNWLLVGSFSPGPFENRMGDVYKCPVDLST 89  
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPPQLLAGCPTVHQCSENTYVKGLC 107

Db 90 ATCEKLNQSTSTIPNVTETKTNMISGLILTRNMGTFGSLTCGLPLWAQCCNQYTTGVC 149  
QY 108 FLFGSNLRQPOKPFALRGCCQEDSDIAFLIDGSGIIPHPFRMKELVSTIMEQLK-- 165  
Db 150 SDISDPF-QLSASFSPATOPCSL-IDVVVCDENSIYPWD--AVXNPLEKFVQGLDIG 205  
QY 166 KXTLPFLMXYSEFRHHTFKBPQNNPRSLIKPITOLLG-RHTATGLRKYVRELFN 224  
Db 206 PTKTQVGLIYANNPRVFNLTNTKTKBEMIVATSTQSYQGDLTNTFGAIQYARKIAYS 265  
QY 225 ITNGARKNAFKILFLITDCEKFGDPLGYEDVTPELDREGVIRY--VLGF--GDAPRSEK 279  
Db 266 AASGGRSRTKVMVVTDESH-DGSMKAVIDOCNWDNIIILFGLAVLYLNALDYKN 324  
QY 280 SRQELNTVASKPPDRHVQFNFAKTLVQNLREKIPAFIEGTQGTGSSSSFEHNSQBGF 339  
Db 325 LIKEIKAIASIPTRYFFNVSDAALLKAGTLGEQIFSGIEGTVGQ-GDNFQWMSQVGF 383  
QY 340 SAALTSNGP--LLSTVGSYDMAGGVPLYTSKEKSTFINMT--RVSDNM-DAYLGVAAAI 394  
Db 384 SADYSQNDILMLGAVGAGWGSTIVQKTSKHGLIFPKQAFQDILQDRNHSSTILGTVAA 443  
QY 395 ILENRVQSLVLAGAPRYOHLGLVAMFRONTGMESNANV-----KGTQIGAFYGASLCSV 448  
Db 444 ISTGESSTHVAGAPRANYTGQIVLYSVN-----ENGNIIVIOAHRGDQIGSYFGVLCVSV 498  
QY 449 DVDSNGSTDLVLICAGHYEOTR--GGQVSVCLPRQQRARWQCDVAVLYGEOGPWGRG 506  
Db 499 DVDKDTITDVLVAGAPMYMSDLKXKEGRVLYFTIKKGLGQHQ---FLEGPGIENITRG 555  
QY 507 AALTLDGVNKGDKLTDVAICAPEEDNRGAVLYFHGTSGSGISPSHSQRIAGS--KLSPR 564  
Db 556 SAIAALSDINDMGFNDVIVGSPLENQNSGAVIYNGHGT-INTKYSQILSDGAFRSH 614  
QY 565 LQYFGOSLSGGDPLWDGLVGLVGAQGHVLLRSQPLVRKVAIMEFNPVARNVFEON 624  
Db 615 LQYFGRSLDGLGDSITDVSIGAFQGVQLWSQSIADVAIEASFTEKI--TLVWKN 672  
QY 625 DOYVXKEAGEVAVCLHVQKSTREDREGIQSWVYTDLALD---SGRPHSAFVNETK 680  
Db 673 AQII-----LKLCP-----SAKFRTKQNNQVAIVITLADGFSRVTSGFLPKENN 721  
QY 681 NSTRRQTVLGLTQC--ETLKLQLPNCIEDPVSPIVLINFLNSLVTGPSLAFGNLRPLVA 738  
Db 722 ERLCQKQVNVVQAQSCPEHIIVIOEPS---DVVNSLDLRAVDISLENPGTS-----PALE 772  
QY 739 EDAQRFLTALPPEKNGNDNICQDDLSITF-----SFMSLDCLVVGCPREFNVTVTRND 794  
Db 773 AVSETAKVFSIPPHKDCGEGDLCSIDLVDVROI PAAQEQPPIVSNQNRILFTSVTLKWK 832  
QY 795 GEDSYRTQVTFPPFLDLVYRKVSTLQNRQSRLWRLACESAST-EVSGALKSTSCSINH 853  
Db 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTEVTCQVAAQSKSVACDVCY 880  
QY 854 PIPPENSEVTNITFDVDSKASIGNKLLKANVTSENMPRTNKTFEQLELPVKYAVYV 913  
Db 881 PALKREQQVTFITNFDNLQ-NLQNAQSLSFQALSESQENKADNLNLIKPLLYDAEI- 938  
QY 914 VTSHGVTKYLNFTASENTSRVMOHOVQVSNLQOR-----SLPISLVFLV----- 958  
Db 939 ---HLRSTNINEFYESSDGNVPSIVHSEFDVCPKFIPLSKVTGTSVPVSMATVHIHIQ 995  
QY 959 -----PVRNLQNTVMDRPOVTF-SENLSTSTCTHKEK 988  
Db 996 YTKENKPNMLYLTQVTDKAGDISCNADINPLKIGQT-----SSSVSPKSENFR---HTKE- 1047  
QY 989 LPSSDPLAELRKAPVNVCSIAVCQRIQCDIPFGIQEENATLKNLSFDWIKYKTSNH 1048  
Db 1048 -----LNCRTACSNTWMLKDVHMKGEYFVNTVTRINNGFASSTFQT 1091  
QY 1049 LLIVSTAEI-LFNDVSFTLLPGQAFVRSQTETKVEFPVNP-----LP--LIVGSSVG 1100  
Db 1092 VQLTAAAEINTYNPEITVI-----EDNTVTIPLMINKPDEKAEVPTGTIIGSIIA 1141

QY 1101 GLLLLALITAAALYKLGPFPRQYKDM 1125  
Db 1142 GILLLALVALVWKLGPFPRKYERM 1166

## RESULT 14

ITAL RAT STANDARD; PRT; 1180 AA.  
ID ITAL RAT  
AC P18614;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)  
DE (CD49a).  
GN ITGAL.  
OS Rattus norvegicus (Rat).  
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90338125; PubMed=2380249;  
RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,  
Esch P., Carbonetto S., Reichardt L.P.;  
RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor  
for laminin and collagen".  
RL J. Cell Biol. 111:709-720(1990).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.  
RX MEDLINE=99313197; PubMed=10386626;  
RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliansky V.,  
Gotwals P.J., Karpusas M.;  
RT "Crystal structure of the alphabeta1 integrin I-domain: insights into  
integrin I-domain function".  
RL FEBS Lett. 452:379-385(1999).  
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-  
E-R IN COLLAGEN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1  
ASSOCIATES WITH BETA-1.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X52140; CAA36384.1; -.  
DR PIR; A35854; A35854.  
DR PDB; 1CK4; 03-MAY-00.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWFA.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
FT SIGNAL; Repeat; Calcium; Magnesium; 3D-structure.  
FT CHAIN 1 28 INTEGRIN ALPHA-1.  
FT DOMAIN 29 1180 EXTRACELLULAR (POTENTIAL).  
FT 29 1142



FT	REPEAT	493	545	FG-GAP 5.
FT	REPEAT	546	605	FG-GAP 6.
FT	REPEAT	608	660	FG-GAP 7.
FT	DOMAIN	1134	1140	POLY-LEU.
FT	CA_BIND	494	502	POTENTIAL.
FT	CA_BIND	558	566	POTENTIAL.
FT	CA_BIND	620	628	POTENTIAL.
FT	DISULFID	76	86	BY SIMILARITY.
FT	DISULFID	666	675	BY SIMILARITY.
FT	DISULFID	681	736	BY SIMILARITY.
FT	DISULFID	789	795	BY SIMILARITY.
FT	CARBOHYD	98	98	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	336	336	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	364	364	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	733	733	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	763	763	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	839	839	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	921	921	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	1011	1011	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	1018	1018	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CARBOHYD	1039	1039	N-LINKED (GLCNAC. .) (POTENTIAL). .
FT	CONFLICT	844	844	I -> L (IN REF. 2).
FT	CONFLICT	909	909	G -> V (IN REF. 2).
FT	CONFLICT	926	926	E -> D (IN REF. 2).
FT	SEQUENCE	1167 AA;	127573 MW;	AE7D3A1C25C1AEAA0 CRC64;

  

Query Match	17.7%;	Score 1042.5;	DB 1;	Length 1167;
Best Local Similarity	28.5%;	Pred. No. 38-62;	Indels 177;	Gaps 44;
Matches 348;	Conservative 201;	Mismatches 495;		

  

QY	1	FNLDTENAMTFOENARG-FGQSVVQLQGRS----	VVVGAPQEIIVAAQNRGSLYQC-----	50
DB	23	FNLDEHHPRLFGPPPEARGFVQLGVGGQRMVLGAPWDGPGDGRGDDVYRCVGGAH	82	
QY	51	-----DYSTG-SCEPIRLQPVVEAVNNSLGLSLAATTSPPQOLLACGPTVHQTCS	99	
DB	83	NAPCAKHLGDLQYLGNSHP-----	ANNHLGMSLLETDDGGFMACAPLWSRACGS	134
QY	100	NTYVXGLCFPLFGSNLRQDPQKPEALRGCPQEDSDIAFLIDGSGSIPIPD-----	PRRMK	154
DB	135	SVFSSGICARVDASFPQGS LAPTAQR-CPDY-MDVVIVLDGNSNIYPWSEVQTELRLRV	192	
QY	155	ELVSTIMEOLKSKTILFSLMQVSEBPRIHFTPKBQNNPNPSRLIKPIYQLLGR-THTAT	213	
DB	193	GKLFIDPEQIQ-----VGLVQVQGESPVHEWSLGDFTKEEVVRAAKNLRRREGRETAK	247	
QY	214	GLRKVVRLEFNTINGARKNAKIFLLTIDGKPF-GDPLGYBVDVIPBLDREGVIRY-VLGF	271	
DB	248	AIMVACTEGFSQSHGGRPEARLLVVYDGEHDSGEEL--PAALKACBAGRVTRYGIAVL	305	
QY	272	GDAPFRSEK-----ROELNTVASKPRDRHVQANNFEALKTQVQUREKI PAIEGTQTGSS	327	
DB	306	CHYLRRQDPSSFLEIRITIASDPPERFFNFVTDAAALTDIVDALGDRIFGLEGSHAENE	365	
QY	328	SSPEHEMSQEGFSAALTSNGPILLSVGSYDWAGGVPLYTSKEKSTPINNTRVDS-----	D	382
DB	366	SSPGLEMSQIGFSTHRLKDGILFGMWGAYDWGGSVLWLEGGHRLFPFPRMALEDEFPALQ	425	
QY	383	MNDAYILGYA-AAIILNRNVSQSLVLGAPRYQHI GLVAMFR-QNTGMWESNANVKGTOIGAY	440	
DB	426	NEAAVIGYSVSSMLLRGGRRLFLSGAPRPHRGKVIAFOLKKDQAVRVAQSLQGEQIGSY	485	
QY	441	FGASICSVDVDNSGSTDLVLCAPHYY--EQTGQGVSVCPILPRGORARWQCDVLYGQE	498	
DB	486	FGSELCPDLTDKGGTDLVLVAAPMFLGPNQKZETGRVVYVLV--GOOSLLTTLGGTLQDPP	543	
QY	499	GQPMWRGFAALTVLGDVNGDKLTDVAIGAPGEDNRGAVILFPHGTSGSGISPSHSQRIAG	558	
DB	544	POD-ARFGFAMGALPDLMQDGFADVAVGAPLEDGHGGLYLYHGTQ-SGVRPHEAQRIAA	601	
QY	559	SKLSRPLQYFGQSLSGQDPLTMDGLVLDLTVGAGHVLLLRSPVLAKVMEPNPREVAR	618	

Db 602 ASMPHALSYFGRSVDGRDLDDLDVAVGAQGAAILSSRPVHLTPSLVTPQAI SV 661  
Qy 619 NVEECNOVVKGEAG--EYEVCLHVOKSTEDRLREGQIQSVVYTDLALDSGRPHSAVP 676  
Db 662 VQDCRR---RGQAVCLTALCFQVTSRTPGRWDH---QFYMFRTASLDEWTAGARAAF 715  
Qy 677 NET--XNSTRQTOVLGHTQCTETLKLQFNCEIDPVSPVILRLNFSLVGTPLSAFGLNR 734  
Db 716 DSGQRLSPRRLSVG-NVCEQLHFPVLD-TSDYLRPVALTVPALDNTTRPG----- 768  
Qy 735 PVLAEDAQRULTALFPPEKQCNENICQDDLSITFSFMSLDC-----LYVGGPREFN 786  
Db 769 PVLNEGSPTSIQKLVPPSKGPGNECVTDLVQ---VNMDIRGSRKAPFVVRGGRKVL 825  
Qy 787 VTVVRNDGSDSYTOVTFPPFLDLSYRKVSTLQONQORSWRLACESASSTEVSGALKS 846  
Db 826 VSTTLNKKENAYNTSLIIFSRNL---HLASLTPQR-ESPIKVECAAPSA-----HA 874  
Qy 847 TSCSINHPIFPENSEVTNITFDVDSKASLG---NKL-----LKANVTSNNMPTNKT 898  
Db 875 RLCSVGHVPVQTGAQVTFLLFEFEFSCSLLSQVFGKLTASSDSLSLERNGTLOENTAOQ--- 931  
Qy 899 EFQLELPVKVAVYVNVTSVSHGVSTKYLNFNTASENTSRVMOHQY-----QVS 943  
Db 932 -----SAYIQVEPH-----LLFSSESTLHYEVHPYGTLPVGPQPEFKTILRVQ 975  
Qy 944 NLG---QRLPISLVFLVP-----VRINQTVIWRDPQVTFSENLSSTCHTKERLPS 991  
Db 976 NLGCYVVSGLIISA--LLPAVAGGNYFLSLSQVI-----TNNASCIVQNLTEPPG 1024  
Qy 992 HSDPLAELRKAPVNVCSIAVCQRIQCDIPFFGIQEEFNATLKGMLSPDVYIKTSHNHLI 1051  
Db 1025 PPVHPELQHTNRLNGSNTQCVVRCHLQGLAKGTEVSVGLLRLVHNEFFRRAKFKSLTV 1084  
Qy 1052 VSTABILFNDVSFTLLPGQGAFVRISQVETKVPPEVFNPLPLIVGSSVGGLLILALITAA 1111  
Db 1085 VSTFELGTEEGSVLQLTEASRWSESLLEV-VQTRPILISLWIMIGSVLGLLMLLIVFC 1143  
Qy 1112 LYKLGPF-----XROYK 1123  
Db 1144 LMKLGPPAHKKIPEEEXREK 1164

Search completed: June 7, 2004, 17:12:42  
Job time : 13.9719 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 39.4889 Seconds  
(without alignments)  
9084.693 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNLDYENAMTFQENARGFGQ.....FKKQYKDMSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:

- 1: sp\_archaea:
- 2: sp\_bacteria:
- 3: sp\_fungi:
- 4: sp\_human:
- 5: sp\_invertebrate:
- 6: sp\_mammal:
- 7: sp\_mhc:
- 8: sp\_organelle:
- 9: sp\_phase:
- 10: sp\_plant:
- 11: sp\_rodent:
- 12: sp\_virus:
- 13: sp\_vertebrate:
- 14: sp\_unclassified:
- 15: sp\_rvirus:
- 16: sp\_bacteriap:
- 17: sp\_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4358	74.1	1151	11 Q9J130	Q9j130 rattus norv
2	3878.5	66.0	1036	11 Q8CA73	Q8ca73 mus musculu
3	3786.5	64.4	920	6 Q28984	Q28984 sus scrofa
4	3457	58.8	1169	4 Q81VA6	Q81va6 homo sapien
5	3297.5	56.1	1169	11 Q9QXH4	Q9qXH4 mus musculu
6	3225.5	54.9	1161	11 Q9QXE7	Q9qe7 rattus norv
7	1512.5	25.7	1161	11 Q9WTV4	Q9wtv4 mus musculu
8	1502	25.5	1160	11 Q9R200	Q9r200 mus musculu
9	1382	23.5	1196	13 Q9R2TF1	Q98tf1 cyprinus ca
10	1341.5	22.8	1086	4 Q96HB1	Q96hb1 homo sapien
11	1329.5	22.6	1187	13 Q98TF0	Q98tf0 cyprinus ca
12	1264.5	21.5	927	6 Q8HZV0	Q8hZv0 bos taurus
13	1142.5	19.4	1167	11 Q88340	Q88340 rattus norv
14	1097	18.7	1167	11 Q88341	Q88341 rattus norv
15	1038.5	17.7	1188	11 Q7TQC3	Q7tc3 mus musculu
16	1034	17.6	1171	13 Q42094	Q42094 gallus gall

#### ALIGNMENTS

##### RESULT 1

Q9J130 ID Q9J130 PRELIMINARY; PRT; 1151 AA.

AC Q9J130;  
DT 01-OCT-2000 (TREMELrel. 15, Created)  
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Integrin beta 2 alpha subunit.  
OS Rattus norvegicus (Rat)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pathallah D.M. Sr., Zerriah K. Jr.;  
RT Cloning of the rat CD11b cDNA sequence."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF268593; AAF81280.1; -  
DR HSSP; P11215; 1BHQ.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00193; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00342; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWFA; 1.  
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.1%; Score 4358; DB 11; Length 1151;

Best Local Similarity 72.3%; Pred. No. 1.2e-314; Indels 2; Gaps 2;

Matches 822; Conservative 151; Mismatches 162;

QY 1 FNLDTENAMTFOENARGFQSVVLOGSRVAVGAPOEIVAAANQSGLYCDDYSTGSCPEI 60  
DB 17 FNLDTENPMTFOENARGFQSVVLOGSRVAVGAPOEIVAAANQSGLYCDDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCPTVHQCSTENTYVKGCLFGLGSLNLRQPOK 120  
DB 77 PLQVPPEAVNMSLGLSLAATTSPPQLLAGCPTVHQCSTENTYVKGCLFGLGSLNLRQPOK 136  
QY 121 FPEALRGCCQEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTLPSLMQYSDEF 180  
DB 137 FPEALRGCCQEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTLPSLMQYSDEF 196  
QY 181 RIHFTFKFQONNPRSLKIPITQLLGRTHRTATGLRQVRELFNITNGARKNAFKILFLL 240  
DB 197 RIHFTFNDPKRNDPKSHVPEIRQLNGRTKTASGIRKVVRELFOKINGARDNAKILVVI 256  
QY 241 TDGEKFGDPLGYEDVITPELDREGVIRYVLGFGDAFRSEKSRQELNLTAVSKPRDHVFOAN 300  
DB 257 TDGEKFGDPLGYEDVITPELDREGVIRYVLGFGDAFRSEKSRQELNLTAVSKPRDHVFOAN 316  
QY 301 NPEALKTQVQUREKI FAIEGTQTSSEFHEMSQEGFSAATISNGPLLSITVGSYDWMAG 360  
DB 317 NPEALKTQVQUREKI FAIEGTQTSSEFHEMSQEGFSAATISNGPLLSITVGSYDWMAG 376  
QY 361 GVFLYTSKSKSTFINNRVDSMDNDAYLVYAAAIILNRVQSVLVLGAPRYCHIGLVAMFR 420  
DB 377 GAFLYPSKXKASFINNRVDSMDNDAYLVYAAAIILNRVQSVLVLGAPRYCHIGLVAMFR 436  
QY 421 QNTGMESANVAGTIGAYFGASLCSVDVDSNGSTDLVLGAPHYEYOTRGQVSVCP 480  
DB 437 QNTGMESANVAGTIGAYFGASLCSVDVDSNGSTDLVLGAPHYEYOTRGQVSVCP 496  
QY 481 PRQGRARWQDAVLGYEQGPWGRFGAALTVLGDVNGDKLTDAI GAPBEENRGAVYLF 540  
DB 497 PRG-RARWQCEALHGDQHPWGRFGAALTVLGDVNGDKLTDAI GAPBEENRGAVYLF 555  
QY 541 HGTSGSGISPHSQRTAGSKLSRLQYEGOSLGGDLTMDGLVTLVGAQGHVILLRQ 600  
DB 556 HGASVASISPHSQRTAGSKLSRLQYEGOSLGGDLTMDGLVTLVGAQGHVILLRQ 615  
QY 601 PVLRLKAIMFNPREVARNYFENCDQVVKGEAGEVRVCLHVOKSTRDRIRREGIOISVVT 660  
DB 616 PVLRLKAIMFNPREVARNYFENCDQVVKGEAGEVRVCLHVOKSTRDRIRREGIOISVVT 675  
QY 661 YDLALDGRPHSRVAVNETKSTRTQVLCILTOTCETLKLQLPNCIEPVPVILRLAF 720  
DB 676 YDLALDGRPHSRVAVNETKSTRTQVLCILTOTCETLKLQLPNCIEPVPVILRLAF 735  
QY 721 SLVGTPLSAFNLPRVLAKDAQLPTALPFPKNCNGNDNICQDDLSITTSFMSLDCLVVG 780  
DB 736 TLVGEPLRSSDLRPLVLAENQRIPTAMFPFKNCGNDNICQDDLSITTSFMSLDCLVVG 795  
QY 781 GPREFNVTVTRNDGDSYQVTPFPPLDLSTRKUSTLQNRQSRWLACASSTEV 840  
DB 796 DSRDFSVTLRNDGDSYQVTPFPPLDLSTRKUSTLQNRQSRWLACASSTEV 854  
QY 841 SGALKSTSCSINHPIPENSEVFNITFDVDSVASLCKNLLKXANTVSENMPRTNKTEF 900  
DB 855 QGVLKSTINHPIPANSEVFNITFDVDSVASLCKNLLKXANTVSENMPRTNKTEF 914  
QY 901 QLELPVAYVMVTSHGVTKYNLFTASNTGRVWQHQQVGNLQORSPLDPSLFLVVP 960  
DB 915 QLELPVAYVMVTSHGVTKYNLFTASNTGRVWQHQQVGNLQORSPLDPSLFLVVP 974  
QY 961 RLQVTLWDRPOVTFSENLSSTCHTKERLPSHSDPLAELAKAPVNCISIAVCCOICDIP 1020  
DB 975 QINKVTWDRPOVTFSENLSSTCHTKERLPSHSDPLAELAKAPVNCISIAVCCOICDIP 1034  
QY 1021 FFGIQEBFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQCGAFVRSQTET 1080  
DB 1035 SFNSKEIFNVTLQNLFLFDWYIETSHDHLVSTAEILFNDVSFTLLPQCGAFVRSQTET 1094  
QY 1081 KYEFPFVNPPLVGVSSVGLLALLIATLALITAGLYKLGFFKQYKDMNSEGPPGAPQ 1137

DB 1095 KYEFPVTVNPPLVGVSSVGLLALLIATLALITAGLYKLGFFKQYKDMNSEGPPGAPQ 1151  
RESULT 2  
Q8CA73 PRELIMINARY; PRT: 1036 AA.  
AC Q8CA73  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Integrin alpha M.  
GN ITGAM OR F730045J24RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573 (2002).  
DR SMEL; AK039444; BAC30350.1; -.  
DR PIR; PT0572; PT0572.  
DR PIR; PT0697; PT0697.  
DR MGD; MGI:96607; Itgam.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; C:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 1.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PRO1185; INTEGRINA.  
DR PRINTS; PRO0453; VWFADOMAIN.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWFA; 1.  
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BFE1E88E77 CRC64;  
Query Match 66.0%; Score 3878.5; DB 11; Length 1036;  
Best Local Similarity 65.6%; Pred. No. 4.7e-279;  
Matches 747; Conservative 126; Mismatches 146; Indels 119; Gaps 2;  
QY 1 FNLDTENAMTFOENARGFQSVVLOGSRVAVGAPOEIVAAANQSGLYCDDYSTGSCPEI 60  
DB 17 FNLDTEHPMTFOENARGFQSVVLOGSRVAVGAPOEIVAAANQSGLYCDDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLAGCPTVHQCSTENTYVKGCLFGLGSLNLRQPOK 120  
DB 77 PLQVPPEAVNMSLGLSLAATTSPPQLLAGCPTVHQCSTENTYVKGCLFGLGSLNLRQPOK 136  
QY 121 FPEALRGCCQEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTLPSLMQYSDEF 180  
DB 137 FPEALRGCCQEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTLPSLMQYSDEF 196  
QY 181 RIHFTFKFQONNPRSLKIPITQLLGRTHRTATGLRQVRELFNITNGARKNAFKILFLL 240  
DB 197 RIHFTFNDPKRNDPKSHVPEIRQLNGRTKTASGIRKVVRELFOKINGARDNAKILVVI 256  
QY 241 TDGEKFGDPLGYEDVITPELDREGVIRYVLGFGDAFRSEKSRQELNLTAVSKPRDHVFOAN 300  
DB 257 TDGEKFGDPLGYEDVITPELDREGVIRYVLGFGDAFRSEKSRQELNLTAVSKPRDHVFOAN 316  
QY 301 NPEALKTQVQUREKI FAIEGTQTSSEFHEMSQEGFSAATISNGPLLSITVGSYDWMAG 360  
DB 317 NPEALKTQVQUREKI FAIEGTQTSSEFHEMSQEGFSAATISNGPLLSITVGSYDWMAG 376

```
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
DB 377 GAFLYTSKDKVFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYVQTRGGQVSVCP 480
DB 437 ENFGWEPHTSIKG----- 450
QY 481 PRQQRARWQCDVLYGEOQQPMGRFGAALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 451 ----- 450
QY 541 HGTSGSGISPSHSORISAGSKLSPRLOYFGOSLGGQDLTMDGLVDTLTVCAOCHVLLBSO 600
DB 451 -----SRIIGAHFSPGLYFGOSLGGQDLTMDGLVDTLTVCAOCHVLLBSO 498
QY 601 PVLVRKALMEFNPREVARNVFCNDQVVGKGEAGEVRVCLHVQKSTRDRRLRGQIQSVYT 660
DB 499 PVLRLBATEFSPKKVARSFACQEQVLKNDKAGEVRVCLVRKNTKORLREGDIQSTWT 558
QY 661 YDLALDSGPHRANVTNSTRQTVLGLTQTCETILKLOLPNCIEDPUSPVLRLNP 720
DB 559 YDLALDPGSRIRAPDETNNTRTQVFLMQKCEITLKLPLPCDDVDSVPIILRLNY 618
QY 721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPFFFEKNCNDNI CODDLSITPSFMSLDCLVVG 780
DB 619 TLVGEPLRSPGNLRPVLAEDAQRLFTALPFFFEKNCNDNI CODDLSITPSFMSLDCLVVG 678
QY 781 GRPEFNVTVRNDGSDSVRTQVTFPPFLDLVSRKVSTI ONORSORSWEL-ACESASSTE 839
DB 679 GPQDFNNSVTLRNDGSDSVRTQVTFPPFLDLVSRKVSTI ONORSORSWEL-ACESASSTE 738
QY 840 VSGALSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTE 899
DB 739 GHGALSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTE 798
QY 900 FOLELPVKAYVMVTSHGVSSTKYNLFTASENTRVMOHQYQVSNLQGSRLPSLVFLVP 959
DB 799 FOLELPVKAYVMVTSHGVSSTKYNLFTASENTRVMOHQYQVSNLQGSRLPSLVFLVP 858
QY 960 VRLNQTVIWRDPQVTSSENLSTCHTKERLPSHSDFLAELKAPVNCVSIANCORIQCDI 1019
DB 859 VQINNVTVMDHPQVIFSQLSSASCHTEQSPHSNFRDQLEBTFVLNCSVAVCKRIQCDL 918
QY 1020 PFFGIEEFNATLKGSLFDWYIKTSHNHLIYSTAEILFNDSVFTLLPQGAQVRSQTE 1079
DB 919 PSFNTQEI FNVTLKGNLSPDWYIKTSHGHLVLSSTELFNDSAPALLPQGSYVRSKTE 978
QY 1080 TKVEPEVPNPPLIVGSSVGGLLLLALITAALYKLGFFKRYKQKDMKSEGGPPGABPQ 1137
DB 979 TKVEPEVHPNPPLIVGSSIGGLVLLALITAGLYKLGFFKRYKQKDMKSEGGPPGABPQ 1036

RESULT 3
QID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Cdlb (Fragment).
GN Cdlb.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
CX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40072; AAB16869.1; -.
DR HSSP; P11215; 1BHQ.
```

```
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007169; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON_TER 1
FT NON_TER 920 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51B350DD5AC CRC64;

Query Match 64.4%; Score 3786.5; DB 6; Length 920;
Best Local Similarity 78.4%; Pred. No. 2.7e-272;
Matches 722; Conservative 85; Mismatches 113; Indels 1; Gaps 1;

QY 118 POKPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMYS 177
DB 1 POKPEALRGCPQSDSDIAFLIDGSGSIINRLDFQMKEFVSTVMQCFQSKTFLPALMYS 60
QY 178 SEFRIHPTFKBFQNNPNSLIKPIITQLLCHTHTATGLRKVRVRELFNITNGARKNAKIL 237
DB 61 EDFYTHPTFNDPKNPSKLLVRPIQLLGRTHRTATGIRKVRVRELPHSKSGARENAKIL 120
QY 238 FLLTDGKFGPLGVEDVPELDREGVIRYVLGPDAPRSKRSQELMTVASKPPDRHVF 297
DB 121 VITDGEKFGPLGVEDVPELDREGVIRYVLGPDAPRSKRSQELMTVASKPPDRHVF 180
QY 298 QANPEALKTVQNLREKIPAIETQTGSSSFHEMSQEGFSAAITNGPILLSTVGSYD 357
DB 181 QVNFEAVKTIQNLQOEXTFAIEGTQGSTSFCEMSQEGFSAAITNGPILLSTVGSYD 240
QY 358 WAGGVLYTSKESKTFINMTVDSDMDAYLGAAAIILNRVQSLVLCAPRYQHIGLV 417
DB 241 WAGGAFLMPKDRVPIINTTRVSDMDAYLGAAAIILNRVQSLVLCAPRYQHIGLV 300
QY 418 MPRONTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYVQTRGGQVSV 477
DB 301 MFKQNSGAWENKADIKSGISQYFGASLCSVDVDSNGSTDLVLGAPHYVQTRGGQVSV 360
QY 478 CPLPRGQARVQCDVLYGEOQQPMGRFGAALTVLGVNGDKLTDVAIGAPGEEDNRGAV 537
DB 361 CPLPGQ-RARQCRVILGEGQHPWSRFGAALTALGVNGDKLTDVAIGAPGEEDNRGAV 419
QY 538 YLFHGTSGSGISPSHSORISAGSKLSPRLOYFGOSLGGQDLTMDGLVDTLTVCAOCHVLL 597
DB 420 YLFHGTSELGSPSHSORISAGSKLSPRLOYFGOSLGGQDLTMDGLVDTLTVCAOCHVLL 479
QY 598 RSQPVLRVKALMEFNPREVARNVFCNDQVVGKGEAGEVRVCLHVQKSTRDRRLRGQIQS 657
DB 480 RSQPVLRVKALMEFNPREVARNVFCNDQVVGKGEAGEVRVCLHVQKSTRDRRLRGQIQS 539
QY 658 VVYDLDLDSGRPHSRVFNENSTRTQTVLGLTQTCETILKLOLPNCIEDPUSPVLRL 717
DB 540 IITDLDLDSGRPHSRVFNENSTRTQTVLGLTQTCETILKLOLPNCIEDPUSPVLRL 599
QY 718 LNFSLVGTPLSAFGNLRPVLAEDAQRLFTALPFFFEKNCNDNI CODDLSITPSFMSLDC 777
DB 600 LNFSLVGTPLSAFGNLRPVLAEDAQRLFTALPFFFEKNCNDNI CODDLSITPSFMSLDC 659
QY 778 VGGPREFNVTVRNDGSDSVRTQVTFPPFLDLVSRKVSTI ONORSORSWEL-ACESAS 837
DB 660 VGGPREFNVTVRNDGSDSVRTQVTFPPFLDLVSRKVSTI ONORSORSWEL-ACESAS 719
QY 838 TEVSGALSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNK 897
DB 720 TEESTALSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNK 779
QY 898 TEFQLELPVKAYVMVTSHGVSSTKYNLFTASENTRVMOHQYQVSNLQGSRLPSLVFL 957
```

```

780 TEFOLELVYAVVYVVVTSLEVTSTKYNFTASEKTRHVIHOYCFNNLGOKLPISVWF 839
958 VVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAAPVNCISAVCORIQ 1017
840 VVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAAPVNCISAVCORIQ 899
1018 DIPFGIOEBFNATLKNLSF 1038
900 DIPFGIOEBFNATLKNLSF 920

RESULT 4
Q81VAG PRELIMINARY; PRT; 1169 AA.
AC Q81VAG;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RC SEQUENCE FROM N.A.
PC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (Sep-2002) to the EMBL/GenBank/DBJ databases.
RE EMBL; BC038237; AAH38237.1; -.
DR GO; GO:0008305; C:intra-cellular complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; WVF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWA; 1.
SQ SEQUENCE 1169 AA; 128521 MW; 178484FEFC79EB6 CRC64;

Query Match
Best Local Similarity 60.78; Pred. No. 1.3e-247;
Matches 685; Conservative 143; Mismatches 295; Indels 6; Gaps 4;

61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGCLFPGSNLRQOPQ 120
80 GLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHVEGRNMYLGLCLFLLGPT--QLTQR 137
121 PPEALRGCPQSDIAFLIDGSGIIIPHDPRMKELVSTIMEQLKSKTLPFLMQYSEEF 180
138 LPVSRQECPRQEQDIVFLIDGSGISSRNTATMNFVRAVVSQRPSTQFSLMQFSNKF 197
181 RIHTEKEFQNNPRLSIKIPITOLGRTHTATGLRKVVRLEFNIWAGKNAFKIPL 240
198 QHTFTFEFRSSNPLSLARSHOGLQFTYATATAIQNVHRLFASVGAARDAAKILVI 257
241 TDGKFGDPLGYEDVPELDRGVIRYVLGFGDAPRSEKSRQELNVTASKPPRDHVPQAN 300
258 TDGKKGSDLOYKDVIPMADAAGIIRYAIGVLAFQNRNSWKLNDIASQSEHIEKVE 317
301 NFEALKTVQNLREKIPAEIGTQTCSSSPHEMSQSGFSAITNSGPLLSTVGSYDWAG 360
318 DFDALKQIQNLREKIPAEIGTQTCSSSPHEMSQSGFSAITNSGPLLSTVGSYDWAG 377

```

```

361 GVFLVTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHILVAMPR 420
378 GAFLYPPNMPPTFINMSQENVMDRDSYLGYSYFELALWKGVSQVSLVIGAPRYQHTGRAVPT 437
421 QNTGHWESNANVKGTCIGAYFGASICSVDVDSNGSTDLVIGAPHYETECCGQVSVCP 480
438 QVSRWRKAEVGTGIGYSYFGASICSVDVDSNGSTDLVIGAPHYETECCGQVSVCP 497
481 PRQARMQCDVLYGEGQPGMRFGAALTVLGDVNGDKLTVAIGAPGEBNARGAVLYF 540
498 PRGWR-RNWCDAVLYGEGQPGMRFGAALTVLGDVNGDKLTVAIGAPGEBNARGAVLYF 556
541 HGTSGSGTSPSHSQRIAGSKLSPRLOYGQSLGSGQDLTMDGLVLTVAQAGHVLRLSQ 600
557 HGVLPESPSHSQRIAGSKLSPRLOYGQSLGSGQDLTMDGLVLTVAQAGHVLRLSQ 616
601 PVLRYKAIMENPREVARNPECNDQVYVKGKAGEVRVCLHVOKSTRDLRLEGQIQSVVT 660
617 PVLWGVSMQFIPABIPRSAFECEQVVSQEQTLVQSNICLYIDKSKNLLSRDLQSSVT 676
661 YDLALDSGRPHRAVFNENSTRQTVGLTQTCETLKLQPCIEDPVPVILRLNF 720
677 LDALDPGRLSPRATFQTKNLSRVRVLGKACENFNLLPSCVEDSVPTITRLNF 736
721 SLVGTPLSAFGNLRPVLAEQAORLFTALPPPEKNCNDNICODDLSITPFSMLDCLVVG 780
737 TLVGRKLLAFNLRPMLAADAQRYFTASLPPEKNCGADHICODNLGISFSPGLKSLVVG 796
781 GPREFNVTYVNDGEDSVTRQVTFPFLDLSYRKVSTLQORSORSWELACESASTEV 840
797 SNLELNAEVWVNDGEDSVTITVTSHPAGLSYRVVAGQKQGLRSLHLTCDSPVGV-- 854
841 SGALKSTCSINHPIPPENSEVFNITFDVDSKASLGKLLKXANVTSENMPRTNKTBF 900
855 SQGTWSTSCRNHLIFRGGAQITFLATDVSPKAVLGDRLLTANVSSENMTPTSKTTF 914
901 QLELPKVAIVVTVSHGVSTKYLNFTAS-ENTSRVMOHQYQVSNLQBSLISLVFLVP 959
915 QLELPKVAIVVTVSHGVSTKYLNFTAS-ENTSRVMOHQYQVSNLQBSLISLVFLVP 974
960 VRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAAPVNCISAVCORIQ 1019
975 VRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAAPVNCISAVCORIQ 1034
1020 PFGIOEBFNATLKNLSFDPWYIKTSHNHLIVSTABILFNDQSVFTLLPQGAFFVRSQTE 1079
1035 PSFSVQEBLDFTLKGNLSFGWVRQILQKKVSVSVVAEITPDTSVYSQLPGEAPMKAQTT 1094
1080 TKVEPEVPEVNPPLVGVSGVGLLALLALITAAALYKLGPEKROYKDMSE 1128
1095 TVLEKTKVENPFLVGVSGVGLLALLALITAAALYKLGPEKROYKDMSE 1143

RESULT 5
Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Leukocyte adhesion glycoprotein p150.95 alpha integrin subunit.
DE ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RT "Isolation of Genes Selectively Expressed by Dendritic Cells."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AF211864; AAF23492.1; -.

```

```

DR HSSP; P11215; 1BHQ.
DR MGP; MG1:96609; C:integrin complex; IEA.
DR GO; GO:0008305; C:integrin receptor activity; IEA.
DR GO; GO:0004895; P:cell-matrix adhesion; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3_A.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWEA; 1.
KW Integrin.
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match 56.1%; Score 3297.5; DB 11; Length 1169;
Best Local Similarity 56.2%; Pred. No. 9.4e-236;
Matches 641; Conservative 174; Mismatches 306; Indels 19; Gaps 7;

Qy 1 FNLDENAMTFOENARFGQSVQLOGSRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 60
Db 20 FNLDAAELTHFMDGAEFGHSLVQYDSVWVVGAPKEIKATNQIGGLYKGYHTGNCEPI 79
Qy 61 RLQVPVAVNMVNLGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLPFLGSLNRQOPQK 120
Db 80 SLQVPVAVNISLGLSLAATNPWILLACGPTVHTCRENIYLTGCLFLSSSPKQS-QN 138
Qy 121 PPEALRGCPQSDIAFLIDGSGIIPHPFRMKELVSTIMQLKSKTFLSLMQVSEEF 180
Db 139 PPTAQCECPKQDQIVELIDGSGISSSTDFPKMLDFVKAMSQLQRPSTFSLMQPBDYF 198
Qy 181 RIHETEFQNNPNSRLIKPITOLLGRTHTATGLRKVVRLENIWNGARKNAFKILFL 240
Db 199 RVHFTFNFNFTSFLSLGSRVQLRGTYTASAKHVIITELFTTQGAQDQATKVLIVI 258
Qy 241 TDGEKFGDPLGYEIVIFELDEGVIRYVLGFGDAPRSEKSRQELNVTASPPRDHVFQAN 300
Db 259 TDGRKQGNLSYDVPVAAEASIIIRVAIGVKAFYNEHQSQELKALASMPSEHYVFSVE 318
Qy 301 NPEALKTVQNLREKIPIAETGTQTSSSSEPHENSQEGFSAITNSGPLLSTGSDYDAG 360
Db 319 NFDALKDIEQLKKEIFAETGTETPSSSTPELNSQEGFSAVFTPDGPFVIGAVGFSWSG 378
Qy 361 GVFLYTSKSKSTFTNTRVDSMDMDAYLGVAALILNRVQSLVGLGAPRYQHIGLVAMPR 420
Db 379 GAFILPSNMRTFTFMNSQENEDMDAYLGYSTALAFKGVHSLILGAPRHQHTGKVIFT 438
Qy 421 QNTGMWESNANKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGGQSVCPPL 480
Db 439 QESRHWPKSEVRGTQIGSVFGASLCSVDMDRSGTDLVIGVPHYHYEHTRGGQSVCPM 498
Qy 481 PRGORAWOCDVLYVGQGPWGFAGALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYTLF 540
Db 499 P-GVGSRWCHCTTLHGQGHFWGFGAALTVLGVDVNGDSLADVAIGAPGEENRGAYIF 557
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSCGQDLFMDGLVDLTGVAQGHVLLRSQ 600
Db 558 HGASRQDIAPSPQRISASQIPSIQYFGQSLSCGQDLTRDGLVDLAVGSKGRVILLRTR 617
Qy 601 PVLKVKALMENPEVARNVPECDQVVKGEAGEVAVCLHVQKSTDRLRREGQIQSVVT 660
Db 618 PILRVSPVTHPTAEIRSRVSECEQVAPBETLSDATVCLIHESPKTLQ--GLRSTVT 675
Qy 661 YDLALDSGRPHSAVFVETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSFVILRLNF 720
Db 676 FDLALDHGLRLSTRAIFXETKTRALTRVKTILGLNKHCSVKLLLPACVEDSVPTILRLNF 735
Qy 721 SLVGTPLSAFGNLRPVLAEDAQRLFTALFPPEKNCMDNITCQDDLSTFTFMSLDCLVWG 780

```

```

Db 736 SLVGVPISSLQNLQPMGLAVDDQTYFTASLPFEKNCQADHICQDDLSVTFGFDLKLTVG 795
Qy 781 GPREFNVTVTRNDGSDSYRTQVTFPFDLDSVRKVSTLQ-----NORSORSWR 829
Db 796 SDLELAVDVTVNSDGDSDYGTVTLEFPYVGLSFRRVABGVFLRKKEQDQWQERGQSHLH 855
Qy 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASIGNKLILKANVTSE 889
Db 856 LACD--STDFRSQGLWSTSCSRHVIFRGSGQMTFLVTFDVSPKAEGLDELLRARVGSB 913
Qy 890 NNMFRNKTETFOLELPVKYAVVNVTVSHGVSTKYLNTASE-NTSRVMQHQYQVSNLQQR 948
Db 914 NNVPGETPKITFOLELPVKYAVVNTMISSHDQFTKYLNFSTSEKETSVEHRRFOVNNLQQR 973
Qy 949 SLPISLAVFLVPVRLNCTVWDREPQVTFSENLSTCTCHERLPSHSDFLAELRKAPVNVCS 1008
Db 974 DVPVSINFVPIELKGEAVW-TVWNSHPQNPFLTCYVRNLKPTQFDLLTFMQSPVLDCS 1032
Qy 1009 IAVCORIQCDIPFGIOBEFNATLKNLSDFMYIKTSHNHLIVSTAELFNDSTVTLPL 1068
Db 1033 IADCLHLRCDIPSGILDELFTLKGNLSFGWISQTLQKKVLLSEAEITFNTSVISQLP 1092
Qy 1069 GQGARVRSOTETKVEPFEVNPPLPLIVSGVGLLILALITRALYKLGREFKROYKMMSE 1128
Db 1093 GQAFLEAQTKTVLEMYKVNVPVPLIVSGVGLLILALITLTKAGFFKROYKEMLEE 1152

RESULT 6
QOQY87 PRELIMINARY; PRT; 1161 AA.
AC QOQY87;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietach G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; EG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWEA; 1.
KW Integrin.
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 54.9%; Score 3225.5; DB 11; Length 1161;
Best Local Similarity 57.2%; Pred. No. 2.1e-230;
Matches 645; Conservative 163; Mismatches 306; Indels 13; Gaps 9;

Qy 2 NLDTENAMTFOENARFGQSVQLOGSRVVGAPQEIIVANQSGSLYQCDYSTGSCPEIR 61

```

Db 21 NLDVEPIVREDAAFGQTWVQGGSRLLVVGAPLEAVAVNQTGRLYDCCAPATGMCQPIV 80  
Qy 62 LOVPVEAVNMSLGLSLAATTPSPOLLACGPTVHTCSENTYVKGCLFLPGSNLRQOPQK 121  
Db 81 LRSPLAVNMSLGLSLVATNNAQLLACGTAORACVKNMYAGSCLLLGSSLL-QRIQAV 139  
Qy 122 PEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTILFSLMOYSEER 181  
Db 140 PASMECPQEQMDIAFLIDGSGSINORDFAQMDKDFKALMGEBFASSTLLFSLMOYSLN 199  
Qy 182 IHFTFKFQNNPNRPSLIKPIITQLLGRTHATGLRKVVRLFNITNGARKNAKILFLLT 241  
Db 200 THFTFTFKNILDPSQLVDPTVOLQGLTYATGRTVMEELFHSKNSRKSANKILLVIT 259  
Qy 242 DGKFGDPLGVEDYVPELDRGVRYVVLGFGDAFRSEKSEKQELMTVASKPPRDRHVFOANN 301  
Db 260 DGQYRPLEYSDVTPADKAGIIRYALGVGDARQEPALKEINTIGSAPPQDHFVKGN 319  
Qy 302 FEALKTQVQUREKIFAIEGTQSGSSSFEHMSQEGFSAITNSGPLLTVGSDYDAGG 361  
Db 320 FAALRSIQRLQEKIFAIEGTQSGSSSFEHMSQEGFSSALTSDGVLGAVGVSFWSGG 379  
Qy 362 VFLATSEKSTFFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFRQ 421  
Db 380 APLPPNTRPTFINMSQENVNDRSYLGYSTAVAFWKGVHSLILGAPRHQHTGKVIFTO 439  
Qy 422 NTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDIWLICAPHYEGTRGGQVSVCP 481  
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRGSDTDLVLICAPHYEGTRGGQVSVFP 499  
Qy 482 FGQRAVOCDAVLGEQOPMGREGAALTVLGDVNGDKLTVAIGAPGEEDNRCAVYLPH 541  
Db 500 -GVRGWCEATLHGEQHPMGREGAALTVLGDVNGDKLTVAIGAPGEEDNRCAVYLPH 558  
Qy 542 GTSGSGISPSHSQRIAGSKLAPRLOQYFQCSLGGQDLTMDGLVLDLTVGAQGHVLLRSOP 601  
Db 559 CASRLIEMPSQSRVTSQSLRLQYFQCSLGGQDLTMDGLVLDLTVGAQGHVLLRSOP 618  
Qy 602 VLKVAIMEPPEPREVARNVPECNDQVWKGXGVRVCLHVKSTRELRGQLOSUVTY 661  
Db 619 LLKVELSRFAPMEVAKAVYQWERTPTVLEAGATVCLTVHKSQPDLL--GNVQGSVRY 676  
Qy 662 DLALDSGPHSRVAFNETKNSRTRQVGLGTQTCETLKLQPNCEIDPSPVILRLNFS 721  
Db 677 DLALDPLGRLISRAIFDEIKNCTLTKRKTGLGDHCEVTKLLPDCVEDAVSPILRLNFS 736  
Qy 722 LVGTPLSAPGNLRLVLAEDAQLFALPFFKXKNCNDNICODDLSITFSNLSLCLVVG 781  
Db 737 LVKDSASP--RNHLPLVAGSQDHITASLFFKXKNCQELLCEGLGISFNFSGLQVLVVG 795  
Qy 782 PREFNVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 841  
Db 796 SPFLTPTVTVNNEGDSYGLVYFAGLSYRVGTQ-QPHQYPLRLACEAPAAQED 854  
Qy 842 GALKSTSCSINHPIIPENSEVTNITFDVDSKASGNKLLKANVTSENKMPRTNKTBP 901  
Db 855 --LRSSSCSINHPIIFREGAKTFTMITFDVSYKAFGLDRLLLAKASSENKMPDNTKTA 912  
Qy 902 LELPKYATVMTVSHGVSTKYLNTASNTSR--VMHQYQVSNLQGRSLPSLVELVPV 960  
Db 913 LELPKYATVMTVSHGVSTKYLNTASNTSR--VMHQYQVSNLQGRSLPSLVELVPV 972  
Qy 961 RLNQTVINDRPPQVTFSEN--TCHTKERLPSHSDPLAELKAPVYVNCISIAVCRIQCDI 1019  
Db 973 LLNGAVMD--VTLSSPAQGVSCVSKMPQNPDPFLTOIQRRSVLDGSIADCLHFRCDI 1029  
Qy 1020 PFGTQEBENATLKNLSFDWIKTSHNHLAVSTABILFNDSVFTLLPQGCAFVRSQTE 1079  
Db 1030 PSLDIQDELDFLIRGLNLSFGWVSQTLQEKVLLVSEABITFDTSVYSQPLGQEFRAQVE 1089  
Qy 1080 TKVEPEVENPLPLIVGSSVGGILLILALITAAVLYKLGFFPKRYQXDM 1126  
Db 1090 TILBEVTVYEPFLVAGSSVGGILLILALITAAVLYKLGFFPKRYQXDM 1136

RESULT 7  
Q9WTV4  
ID Q9WTV4 PRELIMINARY; PRT: 1161 AA.  
AC Q9WTV4;  
DT 01-NOV-1999 (TREMELrel. 12, Created)  
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
DE Integrin alpha L.  
GN ITGAL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBA/2J; TISSUE=Spleen;  
RA Ma R.Z., Teuscher C.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF065902; AAD25885.1; -.  
DR HSSP; P20701; ILFA.  
DR MGD; MGI:96606; Itgal.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR00413; Integrin\_alpha.  
DR InterPro; IPR002035; VWP\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; vwa; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209E431 CRC64;

Query Match 25.7%; Score 1512.5; DB 11; Length 1161;  
Best Local Similarity 33.7%; Pred. No. 5.4e-103;  
Matches 394; Conservative 219; Mismatches 464; Indels 91; Gaps 37;  
Qy 1 FNLDTENAMTFQENA-RGFOQSVVQLGSRVVGAPQEIIVAAQNRGSLYQCDYSTGCEP 59  
Db 24 YLNDTRPTQSLAQAGRHFGYVQLIBDG-VVVGAPGE--GNDTGGLYHRTSSSEFCQP 79  
Qy 60 IRLOVPVEAVNMSLGLSLAATTPSPOLLACGPTVHTCSENTYVKGCLFLPGSNLRQOPQ 119  
Db 80 VSLH-GSNHTSKYLGMTLATDAAGSLACDGLSRTCDQNTYLSGLCYLFPQSLGPM 138  
Qy 120 KFEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTILFSLMOYSE 179  
Db 139 QNRPAVQECMKGVLDVLFVDFGSGSLDRKDFEKLFPMDVMRKLSNTSYQFAAVQFSTD 198  
Qy 180 PRIHPTKEP-QNNPNRPSLIKPIITQLLGRTHATGLRKVVRLFNITNGARKNAKILF 238  
Db 199 CRTEFTLDYVQKKNPDVLLGVSQPMFLNTFRALNVVAVFVKESGAPRPAKVLV 258  
Qy 239 LLVTDGEKF--GDPLGYEDVPELDRGVRYVVLGFGDAFRSEKSEKQELMTVASKPPR 296  
Db 259 IITDGEASDKGNISAAHD-----ITVILGIGKHFFSVQKQKTLHIPASEPVEBF 309  
Qy 297 PQANNFALKTVQNLREKIPATEGTQSGSSSFEHMSQEGFSAITNSGPLLTVGSDY 356  
Db 310 KILDTFEKLDLFDLQRIYAEHTNRQDLTSFNMSLSSSGISADSKGHAVVGAK 369  
Qy 357 DWAGGVF-LVTSKESKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOH 414  
Db 370 DWAGGFLDRDLQGAIFVGOEPLTSDVRGGLYGTVAWMTSRSSRPLAAGADRYQHV 429  
Qy 415 LVAMPR--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYEG 472



Db 780 SPARS-----GPIRLMSASLAVENWTLNSGEDAYVWRLDLPFRGLSFRKVENLQ- 831

Qy 822 QRSQSRWRLACASSTVSGAL-KSTSCSINHPIPPENSEVFNITFDVDSKASLGNKL 880

Db 832 --PHSRMPVSCBELL--TEGSSLLTKLKNVSPPIKAGQVSLQWMTLLSSWEDFV 887

Qy 881 LLKANTISEN-NMPTNKTFQLELPVKVAVVWVTSKYSTKYLNFASENTSRVMOHQ 939

Db 888 ELNGTVHCENENSLOEDNSAATHIPVLPVNLTKEQENSTLIYISFTPKGPKTQOVQHV 947

Qy 940 YQVSNLQGRSLPISLVFLVRLNQTINDRPQ-----VTPSENLS-----TCHKE-RLP 990

Db 948 YQV-----RIQSAYDHNMT-LEALGVPRPHSEDLITVWSVQTDPLVTCHEDLKRP 1001

Qy 991 SHSDPLAKURKAPVNCSTAVCORIQDIPFFGQBEFNATLKNLSFDWYIKTSENHLL 1050

Db 1002 S-----SBAEPLPGV--OPRCPIVF--RWEILQVGTVELSKERKAS-STUS 1045

Qy 1051 IVSTABEILFNDVFTLLPGGAGVRSQTEKVPFFVNPPLPLIVGSSVGGLLLLALITA 1110

Db 1046 LCSSLVSFNSKHFHLYGSKA-SEAQLVKVLDLHEKEMLVHYVLSGIGGLVLLFL 1104

Qy 1111 ALYKLGFFKQYKDM-SEGPGGAEP 1136

Db 1105 ALYKVGFFKRLKXERKOEADGVGPGSP 1131

RESULT 9

Q98TF1 PRELIMINARY; PRT; 1196 AA.

AC Q98TF1;

DT 01-JUN-2001 (T-EMBLrel. 17, Created)

DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE CD11-1.

GN C1A1.

OS *Cyprinus carpio* (Common carp).

OC Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Cyprinus.

OX NCBI\_TaxID=7962;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Peritoneal exudate cells;

RA Kimura M., Nakao M., Miura C., Fujiki K., Yano T.;

RT "Molecular cloning of a leukocyte integrin from the common carp.";

RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB048536; BAB39134.1; --

DR HSP; P20701; 1LPA.

DR GO; GO:0008305; C:integrin complex; IEA.

DR GO; GO:0004190; P:aspartic-type endopeptidase activity; IEA.

DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.

DR GO; GO:0007160; P:cell-matrix adhesion; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001969; Asparticase AS.

DR InterPro; IPR004113; Integrin\_alpha.

DR InterPro; IPR002035; VWFA.

DR Pfam; PF01839; PG-GAP; 3.

DR Pfam; PF00357; integrin\_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int\_alpha; 5.

DR SMART; SM00327; VWFA; 1.

DR PROSITE; PS00141; ASP\_PROTEASE; 1.

DR PROSITE; PS00234; VWFA; 1.

SQ SEQUENCE 1196 AA; 132471 MW; 9369C60787DCA53B CRC64;

Query Match 23.6%; Score 1382; DB 13; Length 1196;

Best local Similarity 31.0%; Pred. No. 2.9e-93;

Matches 369; Conservative 226; Mismatches 451; Indels 144; Gaps 39;

Qy 1 FNLDTENAMTFQENARG-FCQSVVQLQ-GSR--VWVGAPQEIIVANQGRSLYQCDYSTGS 56

Db 32 FNIDTEHFLRENGABEDFFGYSVYQTEFGRKRIIIVGAPLE---GNSTGEMISCTADLQS 88

Qy 57 CEPTEILOVP----VAVVMSLGLSLAATSPQLLACGFTVHOTCSENTYVKGCLFLPQS 112

Db 89 CQ--RLQRPGESVFPFGMSAAVSSAALTS-----CSPYPHECDGNSYLVGVCYQFSS 140

Qy 113 NLQPOQKPEALRCPQEDSDIALIDSGSIIIPHDFRMKELVSTIMEQLKKKTKLPS 172

Db 141 SL-QAVSNFTAYQCSKREVNLFVFDGSSMKTVDFEMNKNFKIDIMKLSNLSIKPA 199

Qy 173 LQVSEERFRIIHTPFKEFQNNPNRSLIKPIITQLLGRTHATGLRKVVRLEFN-ITNGARK 231

Db 200 AVQFSTDVTVDFDNDYQSGSAEELKKE-THMKSILTNHKAIDVILKNLLNSMLSGADS 258

Qy 232 NAFKILFLLTGEKCDPLGYED--VIPELDEGVIRVVLGFGDAFRSEKSOELNVTAS 289

Db 259 KAQKALVITD---GDFSDNDYVNLKCKDEQNTILRYIIVG-----KVDLIEUTQLAS 309

Qy 290 KPRDRHVQANNFEALKTQVQLREKIPAEIGTOTGSSSSFEHEMSQRFSAIITSNQPL 349

Db 310 EFKRNTFVINDYSGKGLDNLQKKIYNIEGSDVAQGRDRQKELSQSGFSVYQBSVI 369

Qy 350 LSTVGSYDWAGVFLYTSKEKSTFIMTRVDSDMN-DAYLGVAALIIILNRVQSLVGNP 408

Db 370 VGSVGSNDWRGALXYTG-SGSDFKETIIPAVNCKDSYMGYSTVLGHRRGVSLFSGAP 428

Qy 409 RYCHIGLVAMPFRONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDLDVLIGAPHYVE 468

Db 429 RAETGLTLVTLTKQNTWVTSNTEGEQISYFGASLLDVSQSDGDFLLVGLAPLFQ 488

Qy 469 QTRGGQSVCLPFRQORARMOQDAVLYGEQ-----GQWGRPGAALTVLGVNGKLTVDV 523

Db 489 SQ-----PRTEGLRVYSLSEQYFQKTLNVQSQTTRGFAASVASIKDLNGDGLSDV 540

Qy 524 ATGAPGSDNRGAVLPHGTSGSLSPSHS-ORLAGSKLSPRLQVFGSGSLGQDLTWDG 582

Db 541 AVGAPLE--NEGVTIYIGDRTHGINBELTQRIISQSVLPGLQFGVSLTQGMNDNDN 598

Qy 583 LVDLVGAQGHVLLRSQVLRVKAIMEFNPREVARNVPECNQDVVKGKEAGEVAVCLHV 642

Db 599 LTDIVIGAQGVILLKARPMVSVAQLSPSKISLNYFECFGS--NAPNAPLTSCTPV 656

Qy 643 QKSTEDRLREGQISV--VTYDIALDSGRPSRAVNETKSTR--RQTQVLGLTQTCET 698

Db 657 TERTSS--TGSLEKLLNVSLNLDVVRGMSRGFFDQSSVSSRTLQSVLLDSSGCCFN 713

Qy 699 LKQLPNCIEDPVSPIVLRNPLSLVGTPLSAFGLNRPVLAEDAQRLFTALPFPFKNCND 758

Db 714 PSIPLMLRCVADTVSPLKIRMNFS--QTEMLS-GNSVAVLDVHSRTEENVVEPFGQNCNSN 770

Qy 759 NTCQDLSITFSFMSLDCLVGGPREFNWTVTVRNDGEDSTRTQVTFPFLDLSYRKVST 818

Db 771 NSCVADLKLNFSTN-NTLVVENQAHTVQVSLANPGGDSYNTSIVLHYPGISLSKFDPA 829

Qy 819 LQNORSQSRWLACESASSTEVSGALKSTCSINHPIPPENSEVFNITFDV--DASKAS 875

Db 830 IKPSR-----SSCGDRDSGATNFTSIDLPVRSCTTQFLGTFRVMKWDNDS 881

Qy 876 LGNKLLKANVTSENMPRTNKTEPQLELPVKIAYVMVTVSHGV-STKILNPTAENTSR 934

Db 882 NRWEIMITANSQNGNM--SDTEVRRESVPQFAVDLAIISLVAEDSVTYMNFSLDROPK 938

Qy 935 VQCHOYQVSNLQGRSLPISLVFLVPVR-----LN 963

Db 939 PLNITYKVENSGKDLFPVSVTLTLTQCPFPVILPHTFSMEVHVSPISSVHQIIMCLLN 998

Qy 964 QTVIWDREPQVTFSENLSSTCHTKERLPKSHSDPLAELKAPVNVNCSIAVCQRIQ----- 1017

Db 999 XHLLFFPSPELSAVQVRSTGWSLREVC-SQF--DLNKSAAVHFNLTADARLQNVKYES 1055

Qy 1018 DIPFGIOEB--FNATLKNLSFDWYIKTSENHLLIVSTAILENDSVFTLLPGQGAQVR 1075

```
Db 1056 KYSFVFPRKDNVPSISAEIYNISLYNQTS-----SELKYNPH-----R 1094
QY 1076 SOTETKVEPPEVNPPLIV-GSSVGGILLALLIITAAALYKLFQKQYKD 1124
Db 1095 SOTETKVE-FVVPSPSLMLIVCTGAVGGFFFLIILFLLKCGFFKRNRPD 1143

RESULT 10
Q96BH1 PRELIMINARY; PRT; 1086 AA.
AC Q96BH1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=lymph;
RA Scrausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008777; AAH08777.1; -.
DR GO; GO:000305; C:integrin complex; IEA.
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00453; VWFADOMAIN.
DR SMART; SK00191; Int_alpha; 5.
DR SMART; SM00327; vwa; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
DR KW Hypothetical protein.
SQ SEQUENCE 1086 AA; 119223 MW; F6FF2546E8C632F9 CRC64;

Query Match 22.8%; Score 1341.5; DB 4; Length 1086;
Best Local Similarity 32.0%; Pred. NO. 2.5e-90;
Matches 374; Conservative 186; Mismatches 434; Indels 175; Gaps 37;

QY 1 FNLDTEANMTFQ--ENARFGOSVVOGSRVVGAPQBIIVANQSGLYQCDYSTGSC 58
Db 26 YNLDVGRASFPSPRAGRHFYRVLVQV-CNGVIVGARGE--GNSGSLYQCGSGTGHCL 81
QY 59 PIRLVQVPEAVNMSIGLSLAATSPQALLACGTVHQTCSNTYVKGCLFPGSNLRQP 118
Db 82 PVTLR-GSNYTSKYLGMTLA--TDP----- 103
QY 119 QKPEALRGCPQSDSIAFLDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSE 178
Db 104 -----TDS----- 117
QY 179 BFAIHFTKFEQNNPHRSLIKPIITOLLGRTHATGLRKVRELFINITGAKRNFKILF 238
Db 118 SYKTEPDFSDYVKKRQDPDALLKHVXEMLLTNTFGAIVYVATEVTFREELGASPDATKVL 177
QY 239 LITDGEKGDPLGYEDVIVPELDRGVIRVYVGFDAFSEKSRQELNTVASKPRDHYVQ 298
Db 178 IITDGE--ATDGNIDAAND-----IIRYIIGIKHFKTQKSEFHLKFAKSPASEFVKI 230
QY 299 ANNFEALKTVQNLREKIFAIBGTQTGSSSFEHMSQEGFSAATISNGPLLSITVGSYDM 358
Db 231 LDTFEKLDLFTLQKKIIVIEGTSKQDLTSFMNELSSSGISADLSRGHAVVAVGAXDW 290
QY 359 AGVFP-LYTSKESKSTINRVDSNDNDAYGVAA--IILNRVQSLVGLGAPRYHIGLY 416
Db 291 AGGFELKADLQDDTFIGNERLTPTEVRAGLYGVYVTLWLPSPROKTSLLASGAPRYOHMGRV 350
```

```
RESULT 11
Q98TF0 PRELIMINARY; PRT; 1187 AA.
AC Q98TF0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CD11-2.
GN CIA2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Peritoneal exudate cells;
RA Kimura M., Fujiki K., Nakao M.;
RT "Molecular cloning of a leukocyte integrin from the common carp."
```

RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AB048537; BAB39135.1; ..	
DR	HSSP; P20701; ILPA	
DR	GO; GO:0008305; C:integrin complex; IEA.	
DR	GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.	
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.	
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.	
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	
DR	InterPro; IPR001969; Aspartate AS.	
DR	InterPro; IPR000413; Integrin_alpha.	
DR	InterPro; IPR002035; VWF_A.	
DR	Pfam; PF01839; FG-GAP; 3.	
DR	Pfam; PF00357; Integrin_A; 1.	
DR	Pfam; PF00092; vwa; 1.	
DR	PRINTS; PR01185; INTEGRINA.	
DR	PRINTS; PR00453; VWFADOMAIN.	
DR	SMART; SM00191; Int_alpha; 5.	
DR	SMART; SM00327; VWA; 1.	
DR	PROSITE; PS00141; ASP_PROTEASE; 1.	
DR	PROSITE; PS00234; VWF_A; 1.	
SQ	SEQUENCE 1187 AA; 131778 MW; 85EDC7CA8B6B1C59 CRC64;	
	Query Match 22.6%; Score 1329.5; DB 13; Length 1187;	
	Best Local Similarity 31.0%; Pred. No. 2.3e-89;	
	Matches 369; Conservative 210; Mismatches 456; Indels 157; Gaps 43;	
QY	1 FNLDTENAMTFQENARG-FGQSVVQLQ-GSR--VVGAPQEIIVAAQNRGLYQCDYSTGS 56	
DB	32 FNIDTEHLEFNGTPEDFGYSYVQTEFGNKRKIIVGAPLE---GNSAGEMYCTADLOS 88	
QY	57 CEPRLQVP----VEAVNMSLGLSLAATSPQQLLAGCPVHQTCSNTVKGKLCFLFGS 112	
DB	89 CK--RLORPQSESVRFPGMSAAVSSAALT-----CSPYFAHECDGNSVLNGVCYQFNS 140	
QY	113 NLRQOPKQFPEARLGCQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMBOLKSKTLFS 172	
DB	141 SL-QAVSNFTAAAYQECSEKREVLNLFVLDGSSSKAVEFDMNKNFIKDVNKKLSNNSIKEA 199	
QY	173 LMOYSEFRTHFTKFFQNNPERSLKPTQLLGRTHATGLRKVVRLFN-ITWGARK 231	
DB	200 AVQFSTIRTVFNDYFQNGASEKLMKE-RHMKSLFTTKALNYLKNLNSVSSGADP 258	
QY	232 NAFKILFLLTGKFGDPLGYED--VPELDRGVIRYVLGPDGAFRSEKSRQELNVTAS 289	
DB	259 NAQKALVIITD---GDPSNDVYNILNICDEQNILRYIIIGV-----KVDLTTLQLAA 309	
QY	290 KPRPDHVFQANNPEALKTQVQNLREKIPAFECTQTGSSSFHEMSQEGESAITSNGPL 349	
DB	310 EPKLNNTFYIOEYNGLGLDLNIOKKIYNIEGSKAHRDROKELSGSGSVVYQESVI 369	
QY	350 LSTVGSYDWAGGVFLYT---SKEKSTFINMTRYDSDMDAYLGYAAAILNRVQSLVLG 406	
DB	370 VGSVGSNDKAGALYEVWVGSGSKFRQTEITDPVAVN---KDSYMGYSTVLGMRHGVSLFSG 426	
QY	407 APYQHIHLVAMFRONTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHY 466	
DB	427 APRAEHTGLVLTFTKNESTVTVMRNTINGEIGSYFGASLLELDVSDGSDFLVLVGAFLF 486	
QY	467 YE-QTRG-GQVSVCPFPRGQARWQCDVAVLGYQ--GQPMGRGAALTVLGVNGDKLTD 522	
DB	487 YQSQPRAEGLYVTL-----SEYQSKTQSTTGRFATSLASLXDLGGLSD 535	
QY	523 VATGAGEEDNRGAVLFIHTSGSGISPSHS-ORLAGSKLSPRLQFGQSLSGQDLTWD 581	
DB	536 VAVGAPLE--NEGVTYILGDGTHGHNPEHAPQIPARSVLPGLQQFGVSLSGQMDMND 593	
QY	582 GLVDLTVGAQGHVLLRSQPLAVKAIMFENPREVARNPECDQVQVKGKEAVEVCLH 641	
DB	594 NLPDIVIGTGGIVLNLNARPMVMSVQLSFNPHEISLNYFECPS--NAPNAELTSCIT 651	
QY	642 VQKSTRDLREGIQSV--VYDIALDSGRPHRAVFNETKNSTR--RQTVLGLTQTCR 697	
DB	652 VTERTS---TGSLEKKNLSMLNLDVVRGMRGPFDPMDSSRTLQQSVLLDSGSGS 708	
QY	698 TLKLOLPNCIEDPVSPVILRLNFS---LVGTPLSAFGLNLRPVLAEADQRLFTALPPPEK 753	
DB	709 NESIFMLRCVADTVSPFKIRMFSTQMLSGNSLAVL-DIQSRTEYVEVL-----FOR 761	
QY	754 NCGNDNICODDLISITFTSMSLDCLVGVGPREPNVTVVRNDEGDSYRTQVTFPPPLDLSY 813	
DB	762 NC-NSNSCVADLKLNFSTN-DTLVVENQAHFTVLVSLANPGDDSYNTSIVLHYPEGLSL 819	
QY	814 RGVSTLQNRQSRQSWRLACESASSFEVSGALKSTSCSINHPIFPENSEVTFNITDV--- 870	
DB	820 SKFDAIKPSRTK-----SCGDRDSGATNRTTCSINLPVRSQTTQFGLTFRVTKW 871	
QY	871 DSKASLGNKLLKANVTSENNMPTNKTEFQLELPKYAVVMVVTSHGV-STKYNLNFAS 929	
DB	872 DYDWDSDRMENTITANSDDNGM---SDMSVRSRIPQFAVELAISLVAEDSVTYLNFSL 928	
QY	930 ENTSRVMOHQYQVSNLQORSPLISLFLVPLVPLNQTIVDRPOVTFSENLSTCHTKERL 989	
DB	929 DRGPRPLNIIYKVNIGLGLPVPVSTLSLPCQ-----TTHVTILTNPFSMQ 974	
QY	990 PSHSDFLAELRAKAPVVC-----SIAVQRIQCDIPFGIOEE 1027	
DB	975 EVHHSFISSYHQ--IIMCLLNKHLFPSPELSAVQTRTGRSLWVC-----VSSISTGEI 1026	
QY	1028 FNATLKGML-----SFDWYIKTSHNLLIVSTAEILFNDSVFTLLPG-QGA 1072	
DB	1027 FRSSV--NLMAEAVLQNVKYESKTSFY-EFRDRHVFNIS-AELNFTSRYNQSTGLKTN 1082	
QY	1073 FVRSQTEKVPFEPFVNPPLPLTVGSSVGGLLLLALITAAALKGLGFFKQYKD 1124	
DB	1083 PHSQTEVKVPEVPIPSRMLIVCTGAVGSGFPFLIIILLKCGFFKMRPD 1134	
RESULT 12		
Q8HZVO	PRELIMINARY; PRT; 927 AA.	
AC	Q8HZVO;	
DT	01-MAR-2003 (TrEMBLrel. 23, Created)	
DT	01-WAR-2003 (TrEMBLrel. 23, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Lymphocyte function-associated antigen 1 (fragment).	
OS	Bos taurus (Bovine).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
OC	Bovidae; Bovinae; Bos.	
OX	NCBI_TaxID=9913;	
RN	[1]_TaxID=9913;	
RP	SEQUENCE FROM N.A.	
RA	Thambikat P., Kannan M.S., Maheswaran S.K.;	
RT	"Sequence of the alpha subunit of bovine lymphocyte function-	
RT	associated antigen 1."	
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF440778; AAN63636.1; ..	
DR	PIR; A32039; A32039.	
DR	GO; GO:0008305; C:integrin complex; IEA.	
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.	
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.	
DR	InterPro; IPR000413; Integrin_alpha.	
DR	InterPro; IPR002035; VWF_A.	
DR	Pfam; PF01839; FG-GAP; 3.	
DR	Pfam; PF00092; vwa; 1.	
DR	PRINTS; PR01185; INTEGRINA.	
DR	SMART; SM00191; Int_alpha; 4.	
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.	
DR	PROSITE; PS00234; VWF_A; 1.	
FT	NON_TER 1	
FT	NON_TER 927	
SQ	SEQUENCE 927 AA; 102523 MW; 02E2CF09917081EC CRC64;	
	Query Match 21.5%; Score 1264; DB 6; Length 927;	
	Best Local Similarity 34.7%; Pred. No. 1.1e-84;	
	Matches 339; Conservative 165; Mismatches 388; Indels 86; Gaps 28;	



DR 489 - EDAFVRIRIEGOMGSGFSGVLCVPTDIDMDGTTDFLLVAAPFTHIRGEGRVYVQVPE 547  
DR 483 GORARMOCDAVLXGEQOPGRFCAALTVLGVNGDKLTDVAICAP-----GEBNNGA 536  
DR 548 - QDASFSLAHTLSGHPCLTNSRFGFAMAAGVDINQKFTDVAICAPLEGFGAGDASGYS 606  
DR 537 VYLPHTGSSGSGISPSHSORLAGSKLSPRLOVFGSLGGDLMDGLVDLTVCAQGHVLL 596  
DR 607 VYIYNGHSG-GLVDSPPQQIRASSVASGLHYFGMSVGGDLFNGDGLADITVSGRSADV 665  
DR 597 LRSQVILRVKAIMFENPREVARNPFCNDQVVGKEAGVRVCLHVQKS---TRDLREG 653  
DR 666 LRSRPVVDLTVSMTFE-----DALPMVFICKM--DVNLCFEVDSSVASEPGLREM 715  
DR 654 QIQSVVTVYDLALDSGRPHSAFVETNSRTRQTVLGLTQTC-----696  
DR 716 FLNFTVDV-----TKQRQLQEDSGSCCLRKNGSGFICEHPWLI 760  
DR 697 ETKLQLPNCIEDPVSIVLRLNPLSLVGTPLSAGNLR-----PVLADAQRLTALP--P 750  
DR 761 STEEL-----CEEDCFSNITIKVYE-----FQTSGGERDYENPTL--DHYKPSAIFQLP 809  
DR 751 FEKCGNDNTQDDLSITTFMSLDCLVYGPRFNTVTVRNGEDSYTQVTFPPPLD 810  
DR 810 YEKDCNKNKVCIAEIQUTN--ISQELVWGVTKVETVNIISLTNSGSDSYMNMALNYPN 868  
DR 811 LSVKRVSTLQWORSQRWELACESASTEVSGALKSTSCSINHPIPPENSEVTFNITFDV 870  
DR 869 LQFKKI-----QKEVSPDQCDPEV---ASVLVKNCKIGHIL--KRSSVNVSVTWQL 918  
DR 871 DSKASLGNKLLKANVTSENNPRTKTEFQLELHPVKVYVMVTVSHGVSTKVLNFTASE 930  
DR 919 EESVFPNRTADITVITISNEKSLARETR---SLQFRHAFIATLSR--PSVMYMN--TSQ 971  
DR 931 NTSRVMOHQVQVNLQORSPLISVLVPLVPLVRLNQTWDRPQVTFSENLST-----CHT 985  
DR 972 SPDSHKFFNVHGENLFAGVFLQICVPIKLQDF-----QIVRVENLTKTDHTECTQ 1025  
DR 986 KERLPSSHDFLAELKAPVNCSTAVCQRIODIPFFGIOEFNATUKMLSPDWYIKTS 1045  
DR 1026 SQEPACGSDPQVQKWHSVVCAI-----TSNKENVTVAABEISVG 1065  
DR 1046 HNHLIIVSTA-----ELFNDVSTPLPGQAFVRSQTETKVEFF---EVPNPLPLIV 1095  
DR 1066 HTQLLRDVSBLPILGEISFNKSLYEGLNAB-----NHRKTIVFLKEETRSPLII 1119  
DR 1096 GSSVGGILLALLALTAALYKLGFFKQYKDMSE 1128  
DR 1120 GSSIGILLVLIIVAILLKQGFKKYQQLNLE 1152  
RESULT 14  
ID C88341 PRELIMINARY; PRT: 1167 AA.  
AC C88341;  
DT 01-NOV-1998 (TReMBLrel. 08, Created)  
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Integrin alpha E2 (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OC NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98056820; PubMed=9394838;  
RA Brenan M., Rees D.J.G.;  
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:  
RT tissue expression reveals phenotypic similarities between  
RT intraepithelial lymphocytes and dendritic cells in lymph.";  
RL Eur. J. Immunol. 27:3076-3079(1997).  
DR EMBL; AF020046; AAC23663.1; --  
DR HSSP; P11215; 1BHQ.

GO: GO:0008305; C:integrin complex; IEA.  
GO: GO:0004895; F:cell adhesion receptor activity; IEA.  
GO: GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; WFAADOMAIN.  
DR SMART; SM00191; Int\_alpha; 4.  
DR SMART; SM00327; vwa; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; WFA; 1.  
FT NON\_TER 1  
SQ SEQUENCE 1167 AA; 128593 MW; E3FD7E18B5CBBEC CRC64;  
Query Match 18.7%; Score 1097; DB 11; Length 1167;  
Best Local Similarity 27.4%; Pred. No. 4.4e-72;  
Matches 330; Conservative 205; Mismatches 427; Indels 244; Gaps 37;  
QY 47 LYQCDYSTGSCBPIRLQVPVEAVNMSGLSLAATSPOLLACGPTVH-----94  
DB 67 LYQC-----ATSIIPDEICQFVVEHILMPKGRYQGV 98  
QY 95 -----QTCSENTVTKGLCFEGNLRQOQPKPPEALRG-----127  
DB 99 LVNHNHNLVLCIQVQSRKPRSLNSELTGACSLTLPNLDLQAAQYFSDLEGLVDLGLSVMS 158  
QY 128 -----CPQED-----SDIAFLIDGSGSIIPHDFRMRKELV 157  
DB 159 GDYSGKSGSTGEBYKSAARLRQAVEEEDDEEAGTEIAVLVDGSGSIPEFQKAKDPI 218  
QY 158 STIMEQL--KSKTTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPITQLLQRTHTATGL 215  
DB 219 STWNRNFKCEKFCNPFALVQYGVQIQTEDPLDLSRDINASLAKVQSVQVKEVTKTASAM 278  
QY 216 RKVURELNTITGARKNPKILFLITDGEKFDQYEDVPELDEGVIRVYVLGFGDAF 275  
DB 279 QHVLDNIFPSRGRKALKWVVLTDGDFRDLPLNTTVISSKXMQGVVRAIGVGNAP 338  
QY 276 RSEKSRQELNTVASKPFRDHPVFOANNFRALKTVQNLREKIPAIESTQTGSSSPHEMS 335  
DB 339 ENNTYRELKLIASDPKAAHTFKVTNYSALDGLLSKLOQRIIHMEGT---VGDTLQYOLA 395  
QY 336 QEGFSAITSNGB-ILSTVGSYDWAGSVFLY--TSKEKSTFINNTRVDSMDNDA---YLGY 390  
DB 396 QTGFSAOILDKGVLLGTGAFNWSGALLINTONGRPLNOT--AKEDFRAAQYSLGY 454  
QY 391 AAAIILRNVRVQSLVGLGAPRYQHIGLVAMFRONTGWESNANVKGTQIGAYFGASLCSVDV 450  
DB 455 SVAAALHKAHGVSYVACAPRHLRGAVFELQKEDGETFMRIEGEQMGSVFGLCPVDI 514  
QY 451 DSNGSTDLVLIGAPHYETRGQGVSVCPPLPRQARWOCDAVLXGEQOPWGRFGAALT 510  
DB 515 NMDGITDFLLVAAPFYHIRGEGRVTVTVV--HQDAPFSLVTVLTSYPLGUTSRFQFAMA 573  
QY 511 VLGDVNGDKLTDVAICAP-----GEBNNGAVYLFHGTSGSGISPSHSQRIAGSKLSPR 564  
DB 574 AVGDINQDKFTDVAICAPLEGFGAGDASGYSVYIYNHSG--GLHASPSQOIRASSVALG 632  
QY 565 LQVFGSLSGQDLTWGLVDLTVGAGVLLRSOPVLRLRQVLRVKAIMEFNREVAENPECN 624  
DB 633 LYTFGMSVSGGLDFSGDDLADITVGSQDVAVVLRSPVVDLTVSMFTFDALP-----685  
QY 625 DQVVKGEAGEVRVCLHVQKS---TRDLREGQIQSVVTVYDLALDSGRPHSAFVETNKN 681  
DB 686 ---MAFKMDVDELCKFVDSAVSPFGLRGMSLNTFTVDV-----TKQ 727  
QY 682 STRRQ-----TQVGLTQCTCTKLQLENCHEDPVSIVLRLNFS 721  
DB 728 KQRLQCADRSCCQCLMKWSGSSSLCEHFGLIETBEL-----CEDDCFSNITIKVSYE 781

```
QY 722 LVGTSLSAFONLAPVLAEDAQRLFTALF--PFRKNCNDNICQDDLSITPSFMSLDCLVV 779
Db 782 F-QTSEERRNHPIL--DHYKESALFQLPYKCKKCKYFCIABIQITTAISQOD-LVV 837
QY 780 GGPREFNVTVNRDGEDSTRTQVTFPPDLDSYRKVSTLQNRORSQSWELACESASTE 839
Db 838 GITKEVTMNISLTSNGSDSYMTNALNYPNLOFKKI-----QKPLSPDIQDDPKPV- 890
QY 840 VSGALSTSCSINHPIPPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRTNKT 899
Db 891 --ASVLVNMCKIGHPII-KRSSVNVSVTWQLESIIPNRTADITVTSNSNEKSLARETH 947
QY 900 FOELPVKYAVYVTVSHGVSTKYLNPTASENTSRVMQHOYQVSNLQORSPLISLVLVP 959
Db 948 --SLOFRHAFIAVLSR--PSVMYMN--TSQSSDHKEFFNFVHGENHFGAVPQLQICVP 1000
QY 960 VRLNQTVINRDPVTFSENLSST-CHTKERLPSHSDFLAELRKAPVNVNSIACVQRIQCD 1018
Db 1001 ITRDL-----QIIRVKHUKTQAH-----ECTOSQPTCGSDPVQNV--- 1039
QY 1019 IPFGIQEFNATL-----KGNLSFDWYIKTSHNHLI--VSTABIL-----FNDSVFTL 1066
Db 1040 -----EWHSVICAITSKENVVAAEISMGHTKQLLRDISELQILGBISFNKSLYEG 1092
QY 1067 LPOGAFVRSQTKVEPP-----EVNPLPLIVGSSVGGILLILALITAAALYKLGPPKROY 1122
Db 1093 LNAB-----NHRKTIVFLKBEKPHSLPLITIGSSIGGLLVVIAILFRCGFFPKKY 1146
QY 1123 KDMWSE 1128
Db 1147 KQNLSE 1152

RESULT 15
Q7TOC3 PRELIMINARY; PRT; 1188 AA.
AC Q7TOC3
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE All integrin.
GN ITGAI1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tiger C.-P., Popova S.N., Velling T., Ekman S., Forsberg E.,
RA Gullberg D.;
RT "allbi integrin is important for mesenchymal cell function:
RT elimination of allbi leads to dwarfism.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Johansson M., Popova S.N.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY124460; AAM62130.1; --
KW Integrin.
SQ SEQUENCE 1188 AA; 133067 MW; 6982BF20D75E4271 CRC64;

Query Match 17.7%; Score 1038.5; DB 11; Length 1188;
Best local similarity 28.1%; Pred. No. 1e-67;
Matches 349; Conservative 182; Mismatches 496; Indels 213; Gaps 46;

QY 1 ENLDTENMTFO-ENARGCQSVVQ--LOGSR-VVVGAPQEIIVANQORSLYCCDYSTGS 56
Db 23 FNMDRNPRVIAGPSAATFGYITVQOHDISGKWLWVGAPMETNMGHOKTGDVYKCPVTQGN 82
QY 57 CEPIL-----QVPEAVNMSIGLSLAATSPOLLACGPTVHQTCSENTYVKGICLFL 110
Db 83 CTKNLGRVTLNVSERKDMRIGLSLATNPXDNSPLACSLWSHCGSSYTTGMSRV 142
```

```
QY 111 GSNLQOQPKPPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMBOLKK--- 166
Db 143 NENFRESKTAP-ALQRC-QTWDIVIVLDGNSIYP-----WVEVQHLINILKIFYIG 195
QY 167 -SKTSLSMQYSEBFRHFTPKFQNNPNRSLIKPIITOLLG-RHTATAGLARKVRELFN 224
Db 196 PQQIQVIGVQYGEDAVHEFLNDYRSVKDVVAAASHIEQRGGTETRTAFGIFPAREAFQ 255
QY 225 ITNGARKNAFKILFLLTDGKEKFGDPLGVEDVPELDRGVIRY---VLFGD--AFRSEK 279
Db 256 --KGRKGKAKVNIIVITDGSHTSP-DIEKVIQSEKQNVTRYAVAVLYGYNRGINPET 312
QY 280 SQEELNVTASKEPPRPHVFOANNFEALKTVQNLREKIFAIEGTOTGSSSSPHEMSQEG 339
Db 313 FLNEIKYIASDDDDKHFFNVYDEALKDIDVALGDRIFSLEGTNK-NETSGLEMSQTGF 371
QY 340 SAAITSNGELSLTVGSYDAGGVFLYTSKEK-----STFINNTRVDSMDNDAYLGVAAAI 394
Db 372 SHVVVEDGILLGAVGAYDMNGAVLAKETISAGKVIPIHRESYLKEPPEELKNHVAIYLGTVTS 431
QY 395 ILNRNV-QSLVILGAPRYQIHGLVAMF-RQNTGMWESNANVKGTQIGAYFGASLCSYDVS 452
Db 432 VVSSRQGVYVAGAPRFXHTGKVLFSMENNRSITIHQALRCEQIGSYFGSITSVDND 491
QY 453 NGSTDLVLIGAPHYYEQTR-GQVSVCELPQGRARWQCDVLYGEOGQWGRFGALTV 511
Db 492 DRVTVLLVGAEMYSEGREKGVVYVNL---RQNRFYVNGTLKDSHSHYQNFARFGSCIAS 548
QY 512 LGDVNGDKLTVAIAPGCEEDNRGAVYLFHGTSGSGISPSHSQRLAGSKLSRLQVFGOS 571
Db 549 VQDLNQDSYNDVVVCGAPLEDSEHGAIYIFHGPO-TNLLKPKMQRITASELAPLOHFGCS 607
QY 572 LSGGQDLTWGLVLTVAQOQHVLRLRPOVLURVKAIMEFNPREVARNVF--ECNDQVVK 629
Db 608 IHGQLDLNEDGLVLAVALGALNAVVMARVQVQINASHLFPEPSKI--NIFHDC----- 659
QY 630 GKEAGEVRVCL-----HVOKSTRDLREGQIQSVTVYDLDLDSGRPHRAVF 676
Db 660 -KRNRDATCLAAFLCFPIFIAPHFQATVG-----IRYNATWDERRYMPRAHL 708
QY 677 NETKNS-TRRQTVLGLTQCTETLKLQPLNCIEDPVSPVILRLNPSLVGTPLSAGNLRP 735
Db 709 DEGGQFTNRAVLSSGQEHQCRINFHVL-DADYVVKPVAFSVEYSLEDP-----DNRP 761
QY 736 VLAEDAQRLLFTALFPPEKNCNDNICQDDL-----SJT 768
Db 762 MLDNGWPTTLVSVSVPFNGCNEDEHCVPELVLDARSDLPTAMEYCORVGRPAQDCSST 821
QY 769 PSFMSLDCLVGGPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRQSRSM 828
Db 822 LSFDTTVFIESTRRVRVAEATLENRGENAYSAVLNISQSENLQF--ASLIQKDDSDNS- 878
QY 829 RLACESASSTVSGALKSTSCSINHPIPFENSEVTNITPDVDSKASLGN-KLLIKANVT 887
Db 879 -IECVNEER-----RLHKVCNVSPYFFRAKAKVAPRLDFPEFSKVFLHLHQIHGAGSD 932
QY 888 SENNMPTNKTETFOLELPVKYAVYVTVSHGVSTKYLAFTASENTS-----RVWQHQQ 939
Db 933 SHEQDSTADDNTALLRFLKYEADVLFTR---SSLSLSEHVKANSLSLSYDGIQPPENCV 989
QY 940 YQVSNLQORSPLISLVL---VPV-----RL-----NOTVIMDRP 971
Db 990 FKVQNLG--FFPIHGVMKKITVPIATRGNNLLMLRDFTDQGNSTCNWGNSTYRSTP 1047
QY 972 QVTFSEMLSSCTKTERLPSSHD---FLAELRKAPVNVNSIACVQRIQCDIPFGIQSEF 1028
Db 1048 T---EEDLS---HAPQNHNSDVVSIIICNRLAP-----SSET 1080
QY 1029 NATLKNLSFDWYIKTSHNHLIIVSTAEI---LFNDSVFTLLPQCGAFVRSQTKVEKVPF 1085
Db 1081 SPYLVGNLWLTSLKALKYRSLEKITVNAALQRFHSPFI-----REEDPSRQVTP 1130
QY 1086 EVPN-----PLPLIVGSSVGGILLILALITAAALYKLGFFK 1119
```

Db 1131 EISKQEDWQVPIWIIIVGSIIGGLLLALLZVLAALWKLGFK 1170

Search completed: June 7, 2004, 17:16:16  
Job time : 46.4889 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:54:39 ; Search time 49.4608 Seconds  
(without alignments)  
6495.175 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTQENARGFGQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq 29Jan04: \*  
1: Geneseqp1980s: \*  
2: Geneseqp1990s: \*  
3: Geneseqp2000s: \*  
4: Geneseqp2001s: \*  
5: Geneseqp2002s: \*  
6: Geneseqp2003as: \*  
7: Geneseqp2003bs: \*  
8: Geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5839	99.3	1153	2 AAW65090	Human Bet
2	5839	99.3	1153	2 AAB07360	Human CD1
3	5839	99.3	1153	5 AAU80252	Human Int
4	5839	99.3	1153	5 ABG61459	Human Bet
5	5839	99.3	1153	5 AA014428	Integrin
6	5839	99.3	1153	7 ADD25615	Binding d
7	5829	99.1	1153	2 AAR04136	Alpha sub
8	3437	58.5	1163	2 AAR07120	p150.95 a
9	3423	58.2	1163	3 AAW65091	Human Bet
10	3423	58.2	1163	3 AAB07361	Human CD1
11	3423	58.2	1163	5 ABG61470	Human Bet
12	3421	58.2	1163	6 ABU07406	Protein d
13	3388	57.6	1161	2 AAR78166	Human bet
14	3388	57.6	1161	2 AAW23049	Human bet
15	3388	57.6	1161	2 AAW57491	Human bet
16	3388	57.6	1161	2 AAW65089	Human bet
17	3388	57.6	1161	2 AAW72825	Human alp
18	3388	57.6	1161	2 AAW73342	Human alp
19	3388	57.6	1161	3 AAB07359	Human alp
20	3388	57.6	1161	5 ABG61468	Human bet
21	3372.5	57.4	1161	2 AAW23064	Human bet
22	3372.5	57.4	1161	2 AAW65106	Human bet
23	3372.5	57.4	1161	2 AAW72837	Human alp
24	3372.5	57.4	1161	2 AAW73343	Human alp
25	3372.5	57.4	1161	3 AAB07376	Human alp

26	3372.5	57.4	1161	5 ABG61485	Human Bet
27	3215.5	54.7	1161	2 AAR78169	Rat alpha
28	3213.5	54.7	1161	2 AAW23062	Rat beta
29	3213.5	54.7	1161	2 AAW60004	Rat alpha
30	3213.5	54.7	1161	2 AAW72824	Rat alpha
31	3213.5	54.7	1161	3 AAB07374	Rat alpha
32	3213.5	54.7	1161	5 ABG61483	Rat Beta2
33	3206.5	54.5	1161	2 AAW65104	Rat beta-
34	3206.5	54.5	1161	2 AAW73345	Rat alpha
35	3201	54.4	1161	2 AAW23061	Mouse bet
36	3201	54.4	1161	2 AAW60003	Mouse alp
37	3201	54.4	1161	2 AAW65103	Mouse bet
38	3201	54.4	1161	2 AAW72836	Mouse alp
39	3201	54.4	1161	2 AAW73347	Mouse alp
40	3201	54.4	1161	3 AAB07373	Mouse alp
41	3201	54.4	1161	5 ABG61482	Mouse Bet
42	3197	54.4	1161	2 AAR78168	Mouse alp
43	3192.5	54.3	1151	2 AAW23059	Rat beta
44	3192.5	54.3	1151	2 AAW60001	Rat alpha
45	3192.5	54.3	1151	2 AAW65101	Rat Beta-

#### ALIGNMENTS

RESULT 1  
AAW65090  
ID AAW65090 standard; protein; 1153 AA.

XX AAW65090;

XX 28-SEP-1998 (first entry)

XX Human Beta-integrin CD11b subunit protein.

Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
lung inflammation; acute respiratory distress syndrome; CD11b subunit;  
rheumatoid arthritis.

XX Homo sapiens.

XX US728533-A.

XX 17-MAR-1998.

XX 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

XX 05-AUG-1994; 94US-00286989.

XX 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

XX Van Der Vieren M, Gallatin WM;

XX WPI; 1998-206565/18.

Screening assay for modulators of integrin binding - using immobilised or  
labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

Example 5; Fig 1A-D; 106pp; English.

This sequence represents a human beta-integrin CD11b subunit which is  
used to describe a method for identifying compounds that modulate the  
interaction of the beta-integrin alpha-d subunit with a binding partner  
of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
-d binding partner, one of which is immobilised and the other of which is  
labelled, in the presence of a test compound, and determining if the  
compound affects binding between the alpha-d polypeptide and alpha-d  
binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
comprising the cytoplasmic, transmembrane or extracellular domain of  
alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
CC and rheumatoid arthritis  
XX  
SQ Sequence 1153 AA;

```
Query Match          99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQGVVQLQGSRRVVGAPQBIIVAAANORGSLYQCDYSTGSCERI 60
DB 17 FNLDTENAMTFQENARGFQGVVQLQGSRRVVGAPQBIIVAAANORGSLYQCDYSTGSCERI 76

QY 61 RLQVPVEAVNNSLGLSLAATTSPPOLLACGPTVHQCSENTYVKGICFLFGSNLRQPOK 120
DB 77 RLQVPVEAVNNSLGLSLAATTSPPOLLACGPTVHQCSENTYVKGICFLFGSNLRQPOK 136

QY 121 FPEALGCPQSDSDIAFLIDGSGIIPDFPRMKELVSTIMEOLKSKTLFSLMOYSSEF 180
DB 137 FPEALGCPQSDSDIAFLIDGSGIIPDFPRMKELVSTIMEOLKSKTLFSLMOYSSEF 196

QY 181 RIHFTKEFQNNPRSLKPIITOLLGRTHATGLRKVVRLEFNIITNGARKNAFKILFL 240
DB 197 RIHFTKEFQNNPRSLKPIITOLLGRTHATGLRKVVRLEFNIITNGARKNAFKILV 256

QY 241 TDCEKGDPLGYEDVPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
DB 257 TDCEKGDPLGYEDVPELDREGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 316

QY 301 NFPAKTVQNLREKIFAIBGTGTGSSSPFHEMSQEGFSAATISNGPLISTVGSYDAG 360
DB 317 NFPAKTVQNLREKIFAIBGTGTGSSSPFHEMSQEGFSAATISNGPLISTVGSYDAG 376

QY 361 GVFLYTSKEKSTINMTRVDSNDAYLGYAAIILNRRVQSLVIGAPRYOHTGLVAMFR 420
DB 377 GVFLYTSKEKSTINMTRVDSNDAYLGYAAIILNRRVQSLVIGAPRYOHTGLVAMFR 436

QY 421 QNTGMWESNANVGTQIGAFYGASLCSVDVDSNGSDLVLIAGAPHYYEOTRGQVSCPL 480
DB 437 QNTGMWESNANVGTQIGAFYGASLCSVDVDSNGSDLVLIAGAPHYYEOTRGQVSCPL 496

QY 481 PRQQRARWQDVLVEYQGGPWRFGAALTVLGVNGDKLTDVAIGAPGEENRGAVYLP 540
DB 497 PRQQRARWQDVLVEYQGGPWRFGAALTVLGVNGDKLTDVAIGAPGEENRGAVYLP 556

QY 541 HGTSGSGISPSHSORLAGSKLSPLOYFGQSLGGQDLTMDGLVLTVGAQGHVLLRQ 600
DB 557 HGTSGSGISPSHSORLAGSKLSPLOYFGQSLGGQDLTMDGLVLTVGAQGHVLLRQ 616

QY 601 PVLVRKAIMEFNPREVARNVFCNDQVYVKGAGEVRVCLVHVKQSTRDLREGQIOSVVT 660
DB 617 PVLVRKAIMEFNPREVARNVFCNDQVYVKGAGEVRVCLVHVKQSTRDLREGQIOSVVT 676

QY 661 YDLALDSGRPHSAVFNKSTRROTQVLGTLCTETLKLQPCNIEDPVPSPVILRLNF 720
DB 677 YDLALDSGRPHSAVFNKSTRROTQVLGTLCTETLKLQPCNIEDPVPSPVILRLNF 736

QY 721 SLVGTPLSAFNLPRVLAEDAQLFTALPPFXKCNQNDNICQDDLSITFSPMSLCLVVG 780
DB 737 SLVGTPLSAFNLPRVLAEDAQLFTALPPFXKCNQNDNICQDDLSITFSPMSLCLVVG 796

QY 781 GPREFNVTVVRNDGDSVETQVTEPEFPLDLSYRKVSTLQNRQSRWLAESASTEV 840
DB 797 GPREFNVTVVRNDGDSVETQVTEPEFPLDLSYRKVSTLQNRQSRWLAESASTEV 856

QY 841 SGALKSTCSINPIPIPNSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 900
DB 857 SGALKSTCSINPIPIPNSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 916

QY 901 QLELPVKYAVYMWVTSKGVSTKYNFTASNTSEVMQHOYQNSLQORSIPSLVPLV 960
DB 917 QLELPVKYAVYMWVTSKGVSTKYNFTASNTSEVMQHOYQNSLQORSIPSLVPLV 976
```

```
QY 961 RLNQTVIMDRPOVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCQRIQCIP 1020
DB 977 RLNQTVIMDRPOVTFSENLSTCHTKERLPSHSDFLAELRKAPVVCNCSIAVCQRIQCIP 1036

QY 1021 FPGIOEFNATLKNLSFDWYIKTSHNHLIIVSTRAEILFNDSVFTLLPQGGAFVRSQET 1080
DB 1037 FPGIOEFNATLKNLSFDWYIKTSHNHLIIVSTRAEILFNDSVFTLLPQGGAFVRSQET 1096

QY 1081 KVEPPEVENPLPLIVGSSVGGLLALITAAALYKLGFFKQYKDMMSGGPGABPQ 1137
DB 1097 KVEPPEVENPLPLIVGSSVGGLLALITAAALYKLGFFKQYKDMMSGGPGABPQ 1153
```

RESULT 2  
AAB07360

ID AAB07360 standard; protein; 1153 AA.

XX AC AAB07360;

XX DT 17-JAN-2001 (first entry)

XX DE Human CD11b protein sequence.

XX KW Human; macrophage infiltration inhibition; alpha d integrin;

XX KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

XX KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

XX KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

XX KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

XX KW rheumatoid arthritis; central nervous system injury; CD11b.

XX OS Homo sapiens.

XX PN WO200029446-A1.

XX PD 25-MAY-2000.

XX PF 16-NOV-1999; 99WO-US027139.

XX PR 16-NOV-1998; 98US-00193043.

XX PR 08-JUL-1999; 99US-00350259.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin MW, Van Der Vieren M;

XX DR WPI; 2000-387751/33.

XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit  
XX macrophage infiltration and reduce inflammation at central nervous system  
XX injury sites.

XX Example 5; Fig 1; 270pp; English.

XX Integrins are a class of membrane-associated molecules that participate  
XX in cellular adhesion. Integrins are made up of an alpha subunit and a  
XX beta subunit. One class of human integrins are restricted to expression  
XX in white blood cells and have a common beta2 subunit: the leukocyte  
XX integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins  
XX have an important role in immune and inflammatory responses. The present  
XX protein sequence is the human integrin alpha subunit CD11b. This sequence  
XX was used in an alignment to identify a novel beta2 integrin alpha  
XX subunit: alpha d (AAB07360 and AAB07359). The present sequence has  
XX approximately 60% identity to the protein sequence of alpha d. The  
XX Alpha\_d gene and protein may be useful in therapy for diseases linked to  
XX aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple  
XX sclerosis, asthma, psoriasis, lung inflammation, acute respiratory  
XX distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency  
XX (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the  
XX inhibition of macrophage infiltration at the site of a central nervous  
XX system injury. The monoclonal antibodies can also be used to detect and  
XX diagnose Crohn's disease

SQ Sequence 1153 AA;			
Query Match 99.1%; Score 5839; DB 3; Length 1153;			
Best Local Similarity 98.9%; Pred. No. 0;			
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;			
QY	1	FNLDTENAMTFOENARGFGQSVVQLGSRVVGAPQEI	VAANQKSLYQCDYSTGSCPEI 60
DB	17	FNLDTENAMTFOENARGFGQSVVQLGSRVVGAPQEI	VAANQKSLYQCDYSTGSCPEI 76
QY	61	RLQVPVEAVNMISGLSLAATTSPQLLACGPTVHQT	CSNTYVKGCLPLFGSNLRQPOK 120
DB	77	RLQVPVEAVNMISGLSLAATTSPQLLACGPTVHQT	CSNTYVKGCLPLFGSNLRQPOK 136
QY	121	FPEALGCGQSDIAFLIDGSGIIPHDFFRMKELVST	IMEQLKSKTLFSLMYSEEF 180
DB	137	FPEALGCGQSDIAFLIDGSGIIPHDFFRMKELVST	IMEQLKSKTLFSLMYSEEF 196
QY	181	RHFTEKEQNPNPESLAKPTOLAGRTHTATGLKRV	RELFNITNGARKNAKFLLEL 240
DB	197	RHFTEKEQNPNPESLAKPTOLAGRTHTATGLKRV	RELFNITNGARKNAKFLLEL 256
QY	241	TGGEKFGDPLGYEDVPELDRGVIRYVILGFGDAP	RSEKSRQELNTVASKPRDHYFOAN 300
DB	257	TGGEKFGDPLGYEDVPELDRGVIRYVILGFGDAP	RSEKSRQELNTVASKPRDHYFOAN 316
QY	301	NPEALKTIVQNLREKIPALEGTQTGSSSFEHEMS	QEGFSAITNSGPLLSTVGSYDMAG 360
DB	317	NPEALKTIVQNLREKIPALEGTQTGSSSFEHEMS	QEGFSAITNSGPLLSTVGSYDMAG 376
QY	361	GVFLYTSKSKSTFINNTRVDSQNDAYLVYAAAI	ILNRNVQSLVLCAPRYQIGLVAMPR 420
DB	377	GVFLYTSKSKSTFINNTRVDSQNDAYLVYAAAI	ILNRNVQSLVLCAPRYQIGLVAMPR 436
QY	421	QNTGMESANVKGTOIGAYFGASLCSVDVDSNGS	TDLVLIGAPHYYEOTRGQSVSCEL 480
DB	437	QNTGMESANVKGTOIGAYFGASLCSVDVDSNGS	TDLVLIGAPHYYEOTRGQSVSCEL 496
QY	481	PRGORARWQCDVLYGEGQGPWCRFGAALTVLGD	VNGDKLTDVAIGAPGEEDNRGAYLIF 540
DB	497	PRGORARWQCDVLYGEGQGPWCRFGAALTVLGD	VNGDKLTDVAIGAPGEEDNRGAYLIF 556
QY	541	HGTSGSGISPSHQSRIAGSKSLPRLQYFGQSLG	SGQDLTMDGLVLTGVAQGEVLLRSQ 600
DB	557	HGTSGSGISPSHQSRIAGSKSLPRLQYFGQSLG	SGQDLTMDGLVLTGVAQGEVLLRSQ 616
QY	601	PVLVRKAIHFNPRVARNVFECNDQVWKGKEAGE	VRVCLVHVKSTRDLREGQIQSVYT 660
DB	617	PVLVRKAIHFNPRVARNVFECNDQVWKGKEAGE	VRVCLVHVKSTRDLREGQIQSVYT 676
QY	661	YDLALDSGRPHSRVAVNETKNSTRROTQVLGLT	QTCETFLKQLPNCIEDPVSPVLRNLF 720
DB	677	YDLALDSGRPHSRVAVNETKNSTRROTQVLGLT	QTCETFLKQLPNCIEDPVSPVLRNLF 736
QY	721	SLVGTPLSAFNLRLPVLADDAQLFTALPFFKNC	QNDNICDDLSITFSFMSLDCLVWG 780
DB	737	SLVGTPLSAFNLRLPVLADDAQLFTALPFFKNC	QNDNICDDLSITFSFMSLDCLVWG 796
QY	781	GRPEFNVTYVRNDGDSVTRQVTFPPFLDLSYR	KVUSTLQNSORSQSWELACESASTEV 840
DB	797	GRPEFNVTYVRNDGDSVTRQVTFPPFLDLSYR	KVUSTLQNSORSQSWELACESASTEV 856
QY	841	SGALKSTCSINHPITFPENSEVTFTITPDVDSK	ASLGNKLLKANVTSENNMPTNKTEF 900
DB	857	SGALKSTCSINHPITFPENSEVTFTITPDVDSK	ASLGNKLLKANVTSENNMPTNKTEF 916
QY	901	QLELPVKYAVYVMTSHGVSSTKYLNETASEN	TSRVWQHGYQVSNLQORSPLSLVFLVPV 960
DB	917	QLELPVKYAVYVMTSHGVSSTKYLNETASEN	TSRVWQHGYQVSNLQORSPLSLVFLVPV 976
QY	961	RLNQTVINDRPQVTFSENLSSTCHTKERLPSHSD	FLAELRKAPVWNCSTAVQRIQCDDIP 1020
DB	977	RLNQTVINDRPQVTFSENLSSTCHTKERLPSHSD	FLAELRKAPVWNCSTAVQRIQCDDIP 1036

QY	1021	FFGIQEEFNATLKGNLSPDWIKTSHNHLIVSTA	EILFNDSVFTLLPGQAGFVRSQDET 1080
DB	1037	FFGIQEEFNATLKGNLSPDWIKTSHNHLIVSTA	EILFNDSVFTLLPGQAGFVRSQDET 1096
QY	1081	KVEPEFVNPPLIVGSSVGGLLILALITAAALYK	LGFPEKQYKXMMSGGPPGABPQ 1137
DB	1097	KVEPEFVNPPLIVGSSVGGLLILALITAAALYK	LGFPEKQYKXMMSGGPPGABPQ 1153
RESULT 3			
AAU80252			
ID	AAU80252 standard; protein; 1153 AA.		
XX			
AC	AAU80252;		
DT	15-JUL-2002 (first entry)		
XX			
DE	Human integrin 1 alpha-M subunit protein.		
XX			
KW	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;		
KW	inflammatory disease; autoimmune disorder; Crohn's disease;		
KW	human immunodeficiency virus; HIV; myocardial infarction;		
KW	Sjorgen's syndrome; rheumatoid arthritis.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	Misc-difference 499..500		
FT	/note= "Encoded by GGG CAG AGG"		
XX	WO200218583-A2.		
PN			
XX			
PD	07-MAR-2002.		
XX			
PF	31-AUG-2001; 2001WO-US027227.		
XX			
PR	01-SEP-2000; 2000US-0229700P.		
XX			
PA	(BLOO-) CENT BLOOD RES INC.		
XX			
PI	Springer TA, Shimoaka M, Lu C;		
XX			
DR	WPI; 2002-382964/41.		
XX			
DR	N-PSDB; ASK50046.		
XX			
PT	Modified integrin-I or integrin I-like domain polypeptide useful as an		
PT	immunogen to produce antibodies specific to polypeptide, comprises a		
PT	disulfide bond such that polypeptide is stabilized in a desired		
PT	conformation.		
XX	Disclosure; page 109-112; 112pp; English.		
XX			
CC	This invention relates to a modified integrin-I or integrin I-like domain		
CC	polypeptide comprising at least one disulfide bond so that the domain is		
CC	stabilised in a desired conformation. The polypeptide of the invention		
CC	may have antiinflammatory or immunosuppressive activities. The		
CC	polypeptides of the invention have an open conformation and are useful as		
CC	immunogens to produce antibodies that selectively bind to integrin I-		
CC	domain, and for identifying a modulator of integrin activity, or of		
CC	interaction of an integrin and a cognate ligand. The polypeptide of the		
CC	invention, or antibodies (preferably anti-LFA-1 antibody) is useful for		
CC	treating or preventing an integrin mediated disorder which is an		
CC	inflammatory or autoimmune disorder in a subject and for inhibiting the		
CC	binding of an integrin to a cognate ligand such as Crohn's disease,		
CC	nephritis; human immunodeficiency virus (HIV), myocardial infarction,		
CC	Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic		
CC	composition comprising the peptide of the invention is useful for		
CC	treating an integrin mediated disorder in a subject. The polypeptides		
CC	and/or active or antigenic fragments are useful as reagents for diagnosis		
CC	of integrin-mediated disorders. The present sequence represents the human		
CC	integrin-I alpha-M protein subunit used to generate the mutant		
CC	polypeptides of the invention		

XX	SQ	Sequence 1153 AA:
		Query Match      99.3%; Score 5839; DB 5; Length 1153;
		Best Local Similarity 98.9%; Pred. No. 0;
		Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY	1	FNLDTENAMTFOENARPGGVSVCLOQSRRVVVGAPQETVAANORGSLVOCDDYSTGSCEPI 60
DB	17	FNLDTENAMTFOENARPGGVSVCLOQSRRVVVGAPQELIIVANORGSLVOCDDYSTGSCEPI 76
QY	61	RLOVPVEAVNMISGLSLAATSPQLLACGPTVHQTCSNTYTVKGLCPLFGSNLRQQPQK 120
DB	77	RLOVPVEAVNMISGLSLAATSPQLLACGPTVHQTCSNTYTVKGLCPLFGSNLRQQPQK 136
QY	121	PPEARLCGPCQEDSDIAFLIDSGSSIIPHDPRRMKELYSTIMEOLKKSKTILFSLMQISEEP 180
DB	137	PPEARLCGPCQEDSDIAFLIDSGSSIIPHDFRMRKEFVSTWMEOLKKSKTILFSLMQISEEP 196
QY	181	RIIHFTFKPEQNPNRSLIXPITOLLGRTHATGLRKVVRELFINITGARKNAPKILFL 240
DB	197	RIIHFTFKPEQNPNRSLVXPITOLLGRTHATGIIRKVVRLEFINITGARKNAPKILVVI 256
QY	241	TDEKEFGDPLGYEUVIPEADREGVIRVYLPGDAFRSEKSRQELINTVASKEPRDHVFQAN 300
DB	257	TDEKEFGDPLGYEUVIPEADREGVIRVYIGVDGAFRSEKSRQELINTASKEPRDHVFQVN 316
QY	301	NFSALKTVQNLREKIIPAIBGTOTGSSSFEHMSQGFSAITSNGPLISTVGSYDWAG 360
DB	317	NFEALKTIQNLRREKIIPAIBGTOTGSSSSFEHMSQBGFSAITSNGPLISTVGSYDWAG 376
QY	361	GVPLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMER 420
DB	377	GVPLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMER 436
QY	421	QNTGMWESNANVKGTQICAYFGASCLSVDVDNSGSTDLVIGAPHYYTEQTRGGQVSCPL 480
DB	437	QNTGMWESNANVKGTQICAYFGASCLSVDVDNSGSTDLVIGAPHYYTEQTRGGQVSCPL 496
QY	481	PRGORARWCDAVLVLYGQGQFWGFFGAALTIVLGDVNGDKLTDVAI GAPGEEDNRGAVTLF 540
DB	497	PRGORARWCDAVLVLYGQGQFWGFFGAALTIVLGDVNGDKLTDVAI GAPGEEDNRGAVTLF 556
QY	541	HGTSGSGISPSHSQRIOIAGSKLSPLRYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 600
DB	557	HGTSGSGISPSHSQRIOIAGSKLSPLRYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 616
QY	601	PVLVRVKAIMEFNPREVARNPFECDNVQVKGKEAGEVRVCLHVQKSTDRLREGQIQSVVT 660
DB	617	PVLVRVKAIMEFNPREVARNPFECDNVQVKGKEAGEVRVCLHVQKSTDRLREGQIQSVVT 676
QY	661	YDLALDSGRPHSRVAFNETKNSTRTOTVIGLTTTCETLKLQLPNCIEDPVSPVILRNLF 720
DB	677	YDLALDSGRPHSRVAFNETKNSTRTOTVIGLTTTCETLKLQLPNCIEDPVSPVILRNLF 736
QY	721	SLVGTPLSAFGNLRPLVAEDAQRFLTALPFPEKNCNDNICODDLSITTFSPMSLDCLVVG 780
DB	737	SLVGTPLSAFGNLRPLVAEDAQRFLTALPFPEKNCNDNICODDLSITTFSPMSLDCLVVG 796
QY	781	GPREFNVTVVRNDGEDSYTQVTFPPPLDLISYRKVSTLONQRSQRWRWLACSSASTEV 840
DB	797	GPREFNVTVVRNDGEDSYTQVTFPPPLDLISYRKVSTLONQRSQRWRWLACSSASTEV 856
QY	841	SGALKSTSCSINHP IIPENSEVTFNIIFPDVDSRASLGNKLLKANVTSENMPRTNKTEF 900
DB	857	SGALKSTSCSINHP IIPENSEVTFNIIFPDVDSRASLGNKLLKANVTSENMPRTNKTEF 916
QY	901	QLELPVKIAYVMVTSIGVSTKILNFITAGSENTGRVMOHQOVSNLGORSIPI SILVFLVPV 960
DB	917	QLELPVKIAYVMVTSIGVSTKILNFITAGSENTGRVMOHQOVSNLGORSIPI SILVFLVPV 976
QY	961	RLNQTVIWDREPQVTFSENLISSTCHTKERLPSHSDFLAE LRKAPVWNGSIAVCQRIQCIDI P 1020

Db	977	RLNQTIVDRPQVTFSENLSSTCHTKERLPESHSDFLAELRKAPVWNCISIAVCQRIQCDIP	1036
Qy	1021	FFGIQEENATLKGNIISFDWYIKTSHNHLIVTAETILFNDVSFTLLPGQGAFYVRSQTET	1080
Db	1037	FFGIQEENATLKGNIISFDWYIKTSHNHLIVTAETILFNDVSFTLLPGQGAFYVRSQTET	1096
Qy	1081	KVPEFVPNPPLIVGSSVGGILLALLITAAALYKLGFFKQYKXDMSEGPPGGAEPQ	1137
Db	1097	KVPEFVPNPPLIVGSSVGGILLALLITAAALYKLGFFKQYKXDMSEGPPGGAEPQ	1153
RESULT 4			
ID	ABG61469		
ID	ABG61469	standard; protein; 1153 AA.	
AC	AC	ABG61469;	
XX	XX		
DT	XX	27-AUG-2002 (first entry)	
DE	XX	Human Beta2 integrin alphaCD11b subunit.	
XX	XX	Beta2 integrin; alpha subunit; CD11c subunit; CD11b subunit; LAD;	
KW	KW	leukocyte adhesion deficiency; inflammatory response; diabetes;	
KW	KW	multiple sclerosis; arthritis; graft atherosclerosis; neuroprotective;	
KW	KW	inflammatory bowel disease; Crohn's disease; ulcerative colitis;	
KW	KW	immune complex alveolitis; leukaemia; ICAM-R; VCAM-1; anti-inflammatory;	
KW	KW	intracellular cell adhesion molecule; vascular cell adhesion molecule;	
KW	KW	locomotor recovery; locomotor damage; locomotor impairment;	
KW	KW	autonomic dysfunction; sensory dysfunction; spinal cord injury.	
OS	XX	Homio sapiens.	
XX	XX		
XX	XX	W02002030980-A2.	
PN	XX		
FD	XX	18-APR-2002.	
XX	XX		
PF	XX	15-OCT-2001; 2001WO-US032059.	
XX	XX		
PR	XX	13-OCT-2000; 2000US-00688307.	
XX	XX		
PA	XX	(ICOS-) ICOS CORP.	
XX	XX		
FI	XX	Gallatin WM, Van Der Vieren M;	
DR	XX	WPI; 2002-463260/49.	
XX	XX		
PT	PT	Use of an anti-alpha-d monoclonal antibodies for promoting locomotor	
PT	PT	recovery, inhibiting locomotor damage, limiting locomotor impairment, or	
PT	PT	limiting autonomic and sensory dysfunction following spinal cord injury.	
XX	XX		
ES	XX	Example 5; Page 191-194; 270pp; English.	
XX	XX		
CC	CC	The invention relates to promoting locomotor recovery, inhibiting	
CC	CC	locomotor damage, limiting locomotor impairment, or limiting autonomic	
CC	CC	and sensory dysfunction following spinal cord injury by administering an	
CC	CC	anti-alpha d (Beta2 integrin alpha2 subunit) monoclonal antibody to a	
CC	CC	spinal cord injury victim. The method also involves the use of a ligand	
CC	CC	selected from ICAM-R or VCAM-1 (intracellular cell adhesion molecule,	
CC	CC	vascular cell adhesion molecule). The method is useful for promoting,	
CC	CC	locomotor recovery, inhibiting locomotor damage, limiting locomotor	
CC	CC	impairment, or limiting autonomic and sensory dysfunction following	
CC	CC	spinal cord injury. In particular, the spinal cord injury comprises	
CC	CC	compression of the spinal cord. The antibodies are also useful for	
CC	CC	reducing inflammation at the site of a central nervous system injury. The	
CC	CC	specification also details the identification of Beta2 integrin alpha d	
CC	CC	cDNAs and proteins, for use in raising the antibodies. Beta2 integrins	
CC	CC	are implicated in diseases such as LAD (leukocyte adhesion deficiency,	
CC	CC	inflammatory response, diabetes, multiple sclerosis, arthritis, graft	
CC	CC	atherosclerosis, inflammatory bowel disease, Crohn's disease, ulcerative	
CC	CC	colitis, immune complex alveolitis and leukaemia. The present sequence is	
CC	CC	a Beta2 integrin alpha subunit sequence included for comparison with the	
XX	XX	Beta2 integrin alpha d protein sequences	
XX	XX		

QY	1021	PFQIOEFNATLKGNSLSPDWIKTSHHLLIVSTABILFNDSVPTLLPGQAFVRSQTEF	1080
DB	1037	PFQIOEFNATLKGNSLSPDWIKTSHHLLIVSTABILFNDSVPTLLPGQAFVRSQTEF	1096
QY	1081	KVEPPEVNPPLIVGSSVGGLLMLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ	1137
DB	1097	KVEPPEVNPPLIVGSSVGGLLMLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ	1153
RESULT 5			
AA014428	standard; protein; 1153 AA.		
ID	AA014428		
XX	AA014428;		
AC	AA014428;		
XX	03-MAY-2002	(first entry)	
DT	03-MAY-2002	(first entry)	
XX	Integrin Mac-1 alpha subunit.		
DE	Integrin Mac-1 alpha subunit.		
XX	Mac-1; integrin alpha subunit; variant integrin inserted domain protein;		
XX	open conformation; integrin related inflammatory disorder;		
XX	integrin related immunological disorder; rheumatoid arthritis; ischaemia;		
XX	reperfusion; hypovolemic shock; infarction; cerebral shock;		
XX	viral infection; cancer; gene therapy; vaccine;		
XX	bioactive agent screening.		
XX	Unidentified.		
OS	Unidentified.		
XX	WO200204521-A2.		
PN	17-JAN-2002.		
PD	17-JAN-2002.		
XX	09-JUL-2001; 2001WO-US021805.		
PF	07-JUL-2000; 2000US-0216600P.		
XX	(CALY ) CALIFORNIA INST OF TECHNOLOGY.		
PR	(BLOO-) CENT BLOOD RES.		
XX	Springer T;		
XX	WPI; 2002-148167/19.		
DR	New integrin I domain protein having alteration in at least 2		
XX	noncontiguous regions and exits in an open conformation, useful for		
PT	treating, preventing or suppressing inflammatory or immunological		
PT	disorders.		
ET	Example 1; Fig 17; 90pp; English.		
XX	The invention comprises structurally biased variant integrin inserted (I)		
XX	domain proteins, wherein the alterations to the protein occur in at least		
CC	two noncontiguous regions. Specifically the variant integrin I domain		
CC	proteins are structurally biased to exist in the open conformation,		
CC	thereby altering the binding ability of the protein. The invention also		
CC	comprises nucleic acids encoding the variant integrin I domain proteins.		
CC	The integrin I domain proteins and nucleic acids are useful for treating,		
CC	preventing or suppressing integrin related inflammatory and immunological		
CC	disorders (e.g. rheumatoid arthritis). The variant integrin I domain		
CC	proteins and nucleic acids can also be used for treating: ischaemia/		
CC	reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral		
CC	infection; and cancer. The variant integrin I domain nucleic acids and		
CC	proteins may be used in gene therapy, as vaccines and to screen for		
CC	bioactive agents. The present amino acid sequence represents the Mac-1		
CC	alpha subunit of integrin		
XX	Sequence 1153 AA;		
SQ	Query Match	99.3%;	Score 5839; DB 5; Length 1153;
	Best local Similarity	98.9%;	Pred. No. 0;
	Matches 1124; Conservative	8; Mismatches	5; Indels 0; Gaps 0;

Query Match 99.3%; Score 5839; DB 5; Length 1153;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARFGQSVVQLQSGSRVVVGAPOEIIAANORGLSYQCDYSTGSCBPI 60  
Db 17 FNLDTENAMTFQENARFGQSVVQLQSGSRVVVGAPOEIIAANORGLSYQCDYSTGSCBPI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGETVHOTCSENTYVKGLCFGLGSLNLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGETVHOTCSENTYVKGLCFGLGSLNLRQOPQK 136  
Qy 121 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKKTLFSLMQVSEBP 180  
Db 137 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKKTLFSLMQVSEBP 196  
Qy 181 RIHFTKEFONNPNRSLKPIITQLLGRTHATGLKVVRELVNININGARKNAFKILFL 240  
Db 197 RIHFTKEFONNPNRSLKPIITQLLGRTHATGLKVVRELVNININGARKNAFKILFL 256  
Qy 241 TDGEKFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVFOAN 300  
Db 257 TDGEKFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVFOAN 316  
Qy 301 NFPAKATVQNLREKIFALEGTCTGSSSPHEHMSQEGSAATSNGLLSTVGSVDWAG 360  
Db 317 NFPAKATVQNLREKIFALEGTCTGSSSPHEHMSQEGSAATSNGLLSTVGSVDWAG 376  
Qy 361 GVFLYTSKEKSTFINNTRVDSNDNDAYLGAAAAIILNRNVQSLVLGAPRYQHIGLVAMPR 420  
Db 377 GVFLYTSKEKSTFINNTRVDSNDNDAYLGAAAAIILNRNVQSLVLGAPRYQHIGLVAMPR 436  
Qy 421 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLI GAPHYVEOTRGQSVCP 480  
Db 437 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLI GAPHYVEOTRGQSVCP 496  
Qy 481 PRQORARWQCDVLYGEGQPMGRFGAALTVDVANGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRQORARWQCDVLYGEGQPMGRFGAALTVDVANGDKLTDVAIGAPGEEDNRGAVYLF 556  
Qy 541 HGTSGSGISPSHQRINGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVILLRSQ 600  
Db 557 HGTSGSGISPSHQRINGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVILLRSQ 616  
Qy 601 PVLKVKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660  
Db 617 PVLKVKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 676  
Qy 661 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQCTETLKLQPNCEIOPVSPDIVRLNF 720  
Db 677 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQCTETLKLQPNCEIOPVSPDIVRLNF 736  
Qy 721 SLVGTPLSAFENLRPVLAEDAQRLEFALPPEKNCNDNIQQDLSITTSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFENLRPVLAEDAQRLEFALPPEKNCNDNIQQDLSITTSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRQSRWLACASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRQSRWLACASSTEV 856  
Qy 841 SGALKSTCSNINPIFPENSEVFNITFDVDSKASLGNKLLKXANTYSENWERTKTEP 900  
Db 857 SGALKSTCSNINPIFPENSEVFNITFDVDSKASLGNKLLKXANTYSENWERTKTEP 916  
Qy 901 QLELPVKYAVVMVTSHGVTSKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVLPVPV 960  
Db 917 QLELPVKYAVVMVTSHGVTSKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVLPVPV 976  
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCISIAVCORICDIP 1020  
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCISIAVCORICDIP 1036  
Qy 1021 FFGIQESFNATLKNLSFDMYIKTSHNHLIVSTAELFNDYSVFTLLPGQGFVRSQTEP 1080  
Db 1037 FFGIQESFNATLKNLSFDMYIKTSHNHLIVSTAELFNDYSVFTLLPGQGFVRSQTEP 1096  
Qy 1081 KVPEFPEVNPPLPLIVGSSVGGLLLLALITAAIYKLGFFKQYKDMMSGGPPGAEPO 1133

Db 1097 KVPEFPEVNPPLPLIVGSSVGGLLLLALITAAIYKLGFFKQYKDMMSGGPPGAEPO 1153  
RESULT 6  
ADD25615  
ID ADD25615 standard; protein; 1153 AA.  
XX  
AC ADD25615;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Binding domain-immunoglobulin fusion protein-associated protein #85.  
XX  
KW Binding domain; immunoglobulin; fusion protein; cytostatic;  
antiarrhythmic; immunosuppressive; antidiabetic; antithyroid;  
neuroprotective; hinge region; immunoglobulin heavy chain;  
CH2 constant region; CH3 constant region; IgG1;  
antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
rheumatoid arthritis; B-cell disorder; melanoma; carcinoma;  
sarcoma;  
type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
XX  
OS Unidentified.  
XX  
PN US2003118592-A1.  
XX  
PD 26-JUN-2003.  
XX  
XX 25-JUL-2002; 2002US-00207655.  
XX  
PR 17-JAN-2001; 2001US-0367358P.  
PR 17-JAN-2002; 2002US-00053530.  
PR 03-JUN-2002; 2002US-0385691P.  
XX  
PA (GENE-) GENE-CRAFT INC.  
XX  
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
XX  
DR WPI; 2003-801317/75.  
XX  
XX New binding domain-immunoglobulin fusion protein, useful for treating a  
subject having or suspected of having a malignant condition or a B-cell  
disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
XX  
PS Disclosure; SEQ ID NO 176; 157pp; English.  
XX  
CC The invention relates to a binding domain-immunoglobulin fusion protein  
comprising a binding domain polypeptide that is fused to an  
immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
CH2 constant region polypeptide that is fused to the hinge region  
polypeptide, and an immunoglobulin heavy chain CH3 constant region  
polypeptide that is fused to the CH2 constant region polypeptide. The  
hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
region polypeptide, derived from (a) having 3 or more cysteine residues;  
where the mutated human IgG1 immunoglobulin hinge region polypeptide  
contains 2 cysteine residues, where the first cysteine is not mutated; a  
mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
(a) having 3 or more cysteine residues, where the mutated human IgG1  
immunoglobulin hinge region polypeptide contains no more than one  
cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
polypeptide, derived from (a) having 3 or more cysteine residues; where  
the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
no cysteine residues. The binding domain-immunoglobulin fusion protein is  
capable of at least one immunological activity comprising antibody  
dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
binding domain polypeptide is capable of specifically binding to an  
antigen. Also included are an isolated polynucleotide encoding the  
binding domain-immunoglobulin fusion protein, a recombinant expression  
construct comprising the polynucleotide operably linked to a promoter,  
a host cell transformed or transfected with a recombinant expression  
construct, producing the binding domain-immunoglobulin fusion protein, a

pharmaceutical composition comprising the binding domain-immunoglobulin fusion protein or polynucleotide and a carrier, and treating a subject having or suspected of having a malignant condition or a B-cell disorder. The binding domain-immunoglobulin fusion protein is useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis, myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple sclerosis or autoimmune disease. The present sequence is a binding domain -immunoglobulin fusion protein-associated protein sequence. Note: The sequence data for this patent formed part of the printed specification and is also available in electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html?DocID=20030118592](http://seqdata.uspto.gov/sequence.html?DocID=20030118592). The authors have not identified the sequences in the printed specification by their SEQ ID number therefore none of the sequences can be explicitly identified.

Query Match	99.3%;	Score 5839;	DB 7;	Length 1153;
Best Local Similarity	98.9%;	Pred. No. 0;		
Matches 1124;	Conservative	8;	Mismatches	5; Indels 0; Gaps 0;
Qy	1	FNLDTENAMTFOENARGFGQSIVOLQGSRVVVGAPQEIVAANRGSLYCCDYSTGSCBPI	60	
Dd	17	FNLDTENAMTFOENARGFGQSIVOLQGSRVVVGAPQEIVAANRGSLYCCDYSTGSCBPI	76	
Qy	61	RLOVPVEAVNMISLGLSAAATTPPOLACAGTVTAQTCSNTYVYRGLCLFSGNLRQQPQK	120	
Dd	77	RLOVPVEAVNMISLGLSAAATTPPOLACAGTVTAQTCSNTYVYRGLCLFSGNLRQQPQK	136	
Qy	121	FPEALRCGPQSDIAFLIDGGSGIIPHDFRMKELVSTMEOUKKSXTLSLMQYSSEFP	180	
Dd	137	FPEALRCGPQSDIAFLIDGGSGIIPHDFRMKEFVTVMEOUKKSXTLSLMQYSSEFP	196	
Qy	181	RIHETEXEKFQNNPNRSILXIPITQLGRTHATGLRKVVURELENTNGARKNAFKILL	240	
Dd	197	RIHETEXEKFQNNPNRSILVKPIITQLGRTHATGIKKVVURELPNLTNGRNKNAFKILVVI	256	
Qy	241	TGSEKFGDPLGYEDVIPBELDEGVIRYYVLGFDAFRSEKSRQELNTVASKPPRHVFQAN	300	
Dd	257	TGSEKFGDPLGYEDVIPPEADREGVIRYVIGVDFAFRSEKSRQELNTIASKPPRHVFQVN	316	
Qy	301	NFEALKTIQNOLREKI FAIEBTQTQGSSSPHEMSQBGSAAITSNGPLLSTVSQSYDWAG	360	
Dd	317	NFEALKTIQNOLREKI FAIEBTQTQGSSSPHEMSQBGSAAITSNGPLLSTVSQSYDWAG	376	
Qy	361	GVFLYTSKEKSTFTINMTRVDSMDNDAYLGYAIIILNRNVQSLVLGAPRYOHIGLVAMFR	420	
Dd	377	GVFLYTSKEKSTFTINMTRVDSMDNDAYLGYAAIIILNRNVQSLVLGAPRYOHIGLVAMFR	436	
Qy	421	QNTGWSSNANVKGTQIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYPTQTGGQSVQCPL	480	
Dd	437	QNTGWSSNANVKGTQIGAYFGASLCSDVDVDSNGSTDVLVIGAPHYPTQTGGQSVQCPL	496	
Qy	481	PRGORARWQCDVALYVGOGQPWGFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYP	540	
Dd	497	PRGORARWQCDVALYVGEGQPWGFRGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYP	556	
Qy	541	HGTSGSGISPSSHQRISKLSPRLQYFGQSLSGQDLTMDBGLVDLTVGAQHVLILLSQ	600	
Dd	557	HGTSGSGISPSSHQRISAGSKLSPRLQYFGQSLSGQDLTMDBGLVDLTVGAQHVLILLSQ	616	
Qy	601	PVLAVKAIMEFNPREVAARNVFECHNDQVWKGEAGEVRVCLHVOKSTRDRLBREGIQSVYT	660	
Dd	617	PVLAVKAIMEFNPREVAARNVFECHNDQVWKGEAGEVRVCLHVOKSTRDRLBREGIQSVYT	676	
Qy	661	YDLALDSGRPHSRVAFNETKNSRRTOTVLGITQTCETLKQLPNCIEDPDVSPVILRLNF	720	
Dd	677	YDLALDSGRPHSRVAFNETKNSRRTOTVLGITQTCETLKQLPNCIEDPDVSPVILRLNF	736	
Qy	721	SLVGTPLSAFGNLRPVTLAEADAQRIFTALPFPEKNCNGNDNI CODDLSITFFSMISLDCLVYG	780	
Dd	737	SLVGTPLSAFGNLRPVTLAEADAQRIFTALPFPEKNCNGNDNI CODDLSITFFSMISLDCLVYG	796	

Qy	781	GPREFNVTVTRNDGSDSYRTQVTPFPPLDLSTRKVTSTLQNQSRQSRWLACESASSTEV	840
Db	797	GPREFNVTVTRNDGSDSYRTQVTPFPPLDLSTRKVTSTLQNQSRQSRWLACESASSTEV	856
Qy	841	SGALKSTCSINHPPIPNSESEVFNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF	900
Db	857	SGALKSTCSINHPPIPNSESEVFNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF	916
Qy	901	QLELPVKYAYVMVTVSHGVSTKYKLNFTASENTSRVVMQHOYQVSNLQORSPLISLVFLVPV	960
Db	917	QLELPVKYAYVMVTVSHGVSTKYKLNFTASENTSRVVMQHOYQVSNLQORSPLISLVFLVPV	976
Qy	961	RLNQTVTWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIOCDIP	1020
Db	977	RLNQTVTWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCORIOCDIP	1036
Qy	1021	FFGIQSEFNATLKGNSLFDWYIKTSHNHLIVTASIELFNDSVFTLLPQCGAFVRSOTET	1080
Db	1037	FFGIQSEFNATLKGNSLFDWYIKTSHNHLIVTASIELFNDSVFTLLPQCGAFVRSOTET	1096
Qy	1081	KVPEPFVENVPLPLIVGSSVCGILLIALLITAAALYKLGFFKQYKDMMSGGPGCAEPQ	1137
Db	1097	KVPEPFVENVPLPLIVGSSVCGILLIALLITAAALYKLGFFKQYKDMMSGGPGCAEPQ	1153
RESULT 7			
RAR04136			
ID	RAR04136 standard; protein; 1153 AA.		
XX	RAR04136;		
AC			
XX			
XX	25-MAR-2003 (revised)		
DT	07-SEP-1990 (first entry)		
XX	Alpha subunit of Mac-1 leukocyte adhesion receptor.		
DE			
XX			
XX	Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;		
KW	non-specific defence system; integrin gene superfamily.		
XX			
OS	Synthetic.		
XX			
FR	Key	Location/Qualifiers	
FT	Region	1..16	
FT	Modified-site	/label= signal peptide	
FT	Modified-site	86..88	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	240..242	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	391..393	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	469..471	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	693..695	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	697..699	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	735..737	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	802..804	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	881..883	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	901..903	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	912..914	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	941..943	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	947..949	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	979..981	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	994..996	

FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1022..1024	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1045..1047	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1051..1053	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1076..1078	
FT	Modified-site	/label= putative N-glycosylation site	
FT	Modified-site	1106..1134	
FT	Region	/label= putative transmembrane region	
XX		EP364690-A.	
XX		25-APR-1990.	
XX		17-AUG-1989;	89EP-00115159.
XX		23-AUG-1988;	89US-00235353.
XX		09-MAR-1989;	89US-00321239.
XX		(DAND ) DANA FARBER CANCER INST INC.	
XX		Springer TA, Corbi A;	
XX		WPI; 1990-125938/17.	
XX		N-ESDB; AAQ04043.	
XX		New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating	
XX		inflammation and viral infections, and in diagnosis.	
XX		Disclosure; Page ?; -pp; English.	
XX		Mac-1 alpha subunit is involved in the response to inflammation, i.e.	
XX		recognition of and migration to sites of inflammation. It also attaches	
XX		to cellular substrates as part of this function making it useful in	
XX		visualising endothelial tissue. Mac-1 is a member of the Integrin Gene	
XX		superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25	
XX		-MAR-2003 to correct PA field.)	
XX		Sequence 1153 AA;	
XX		Query Match	99.1%; Score 5829; DB 2; Length 1153;
XX		Best Local Similarity	98.8%; Pred. No. 0;
XX		Matches 1123; Conservative	8; Mismatches 5; Indels 0; Gaps 0;
QY		1 FNLDTENAMTQENARGGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60
DB		17 FNLDTENAMTQENARGGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	76
QY		61 RLQVPVEAVNLSGLSLAATSPPLLACGPTVHQTCSNTYVVKGLCPFGSNLRQOPQK	120
DB		77 RLQVPVEAVNLSGLSLAATSPPLLACGPTVHQTCSNTYVVKGLCPFGSNLRQOPQK	136
QY		121 PPEALRGCPQSDSDIAFLIDSGSTIIPHDPRMKELVSTIMEQLKSKTLPFLMQYSEEF	180
DB		137 PPEALRGCPQSDSDIAFLIDSGSTIIPHDPRMKELVSTIMEQLKSKTLPFLMQYSEEF	196
QY		181 RIHPTFKFQNNPNSRLIKPTTOLLGTHTATGLKVVRELFNITNGARKNAKILFL	240
DB		197 RIHPTFKFQNNPNSRLIKPTTOLLGTHTATGLKVVRELFNITNGARKNAKILVVI	256
QY		241 TDGEXFGDPLGYEDVPEADREGVIRYVVGDAFPESEKSRQELNTVASKPRDHFQAN	300
DB		257 TDGEXFGDPLGYEDVPEADREGVIRYVVGDAFPESEKSRQELNTVASKPRDHFQAN	316
QY		301 NFEALTKVQNLREKIPALEGTQTCSSSFHEMSQEGFSAATNSGILLSTVGSYDWAG	360
DB		317 NFEALTKVQNLREKIPALEGTQTCSSSFHEMSQEGFSAATNSGILLSTVGSYDWAG	376
QY		361 GVFLYTSKEKSTFFINWTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR	420
DB		377 GVFLYTSKEKSTFFINWTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR	436

QY	421	QNTGMWESNANVKCTQIGAYFGASICSVDVDSNGSTDLVLIGAPHYVYQTRGQSVVCP	480
DB	437	QNTGMWESNANVKCTQIGAYFGASICSVDVDSNGSTDLVLIGAPHYVYQTRGQSVVCP	496
QY	481	PRQORARWQCDVAVLYGEQCPWGRFGAALTVDVGVNGDKLTVDVAGAPGEEDNRGAVYLF	540
DB	497	PRQORARWQCDVAVLYGEQCPWGRFGAALTVDVGVNGDKLTVDVAGAPGEEDNRGAVYLF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSISSGGQDLTMDGLVDLTVCAQGHVLLRSQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSISSGGQDLTMDGLVDLTVCAQGHVLLRSQ	616
QY	601	PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRLBQSIQSVVT	660
DB	617	PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRLBQSIQSVVT	676
QY	661	YDLALDSGRPHSRVAVNETKNSRQTQVGLTQTCETLKLQLPNCIEBPVPIVRLNF	720
DB	677	YDLALDSGRPHSRVAVNETKNSRQTQVGLTQTCETLKLQLPNCIEBPVPIVRLNF	736
QY	721	SLVGTPLSAFGNLRPVLAEDAQLFTALFPFPEKNCNDNICODDLSITFSFMSLDCIIVG	780
DB	737	SLVGTPLSAFGNLRPVLAEDAQLFTALFPFPEKNCNDNICODDLSITFSFMSLDCIIVG	796
QY	781	GPREFNVTVVRNDGEDSYRTQVTPFPPLDLSYRKVSTLONQBSQRSWRLACESASSTEV	840
DB	797	GPREFNVTVVRNDGEDSYRTQVTPFPPLDLSYRKVSTLONQBSQRSWRLACESASSTEV	856
QY	841	SGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTEF	900
DB	857	SGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPRTNKTEF	916
QY	901	QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVLPV	960
DB	917	QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVLPV	976
QY	961	RLNOTVINDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQDIP	1020
DB	977	RLNOTVINDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQDIP	1036
QY	1021	PFQIQEFPNATLKNLSFDWYIKTNSHLLIVSTABILFNDVSFTLLPGQAFVRSQTET	1080
DB	1037	PFQIQEFPNATLKNLSFDWYIKTNSHLLIVSTABILFNDVSFTLLPGQAFVRSQTET	1096
QY	1081	KVEPFEVNPPLPIVGVSSVGGILLALITAAALYKLGPFKXROKDMSEGGPPCAEPQ	1137
DB	1097	KVEPFEVNPPLPIVGVSSVGGILLALITAAALYKLGPFKXROKDMSEGGPPCAEPQ	1153
XX		RESULT 8	
XX		AAR07120	
XX		ID AAR07120 standard; protein; 1163 AA.	
XX		AC AAR07120;	
XX		25-MAR-2003 (revised)	
XX		05-FEB-1991 (first entry)	
XX		p150.95 alpha subunit encoded by clone lambdaX47.	
XX		p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;	
XX		rhinovirus.	
XX		Synthetic.	
XX		Key	Location/Qualifiers
XX		Region	1..19
XX		Region	/label= signal peptide
XX		Region	20..44
XX		Modified-site	/label= N-terminus
XX		Modified-site	61..63
XX		Modified-site	/label= glycosylation site



KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
 KW type-1 diabetes; atherosclerosis; multiple sclerosis; asthma;  
 KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;  
 KW rheumatoid arthritis.  
 XX Homo sapiens.  
 XX US5720533-A.  
 XX 17-MAR-1998.  
 XX 07-JUN-1995; 95US-00485618.  
 XX 23-DEC-1993; 93US-00173497.  
 XX 05-AUG-1994; 94US-00286889.  
 XX 21-DEC-1994; 94US-00362652.  
 XX (ICOS-) ICOS CORP.  
 XX Van Der Vieren M, Gallatin WM;  
 XX W21; 1998-206565/18.  
 XX Screening assay for modulators of integrin binding - using immobilised or  
 XX labelled alpha-d polypeptide, useful for, e.g. treating type-1 diabetes.  
 XX Example 5; Fig 1A-D; 106pp; English.  
 XX This sequence represents a human beta-integrin CD11c subunit which is  
 XX used to describe a method for identifying compounds that modulate the  
 XX interaction of the beta-integrin alpha-d subunit with a binding partner  
 XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
 XX -d binding partner, one of which is immobilised and the other of which is  
 XX labelled, in the presence of a test compound, and determining if the  
 XX compound affects binding between the alpha-d polypeptide and alpha-d  
 XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
 XX comprising the cytoplasmic, transmembrane or extracellular domain of  
 XX alpha-d. Compounds that modulate alpha-d binding could be used to treat  
 XX diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis,  
 XX asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
 XX and rheumatoid arthritis  
 XX Sequence 1163 AA;  
 Query Match 58.2%; Score 3423; DB 2; Length 1163;  
 Best Local Similarity 60.2%; Pred. No. 2.7e-274;  
 Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;  
 1 FNLDTNMTFQNGARGQGVVQLQGSVVVVGAPQEIIVAAQNGSLVQCDYSTGSCPEI 60  
 20 FNLDTNMTFQNGARGQGVVQLQGSVVVVGAPQEIIVAAQNGSLVQCDYSTGSCPEI 79  
 61 RLQVPEAVNMSLGLSLATTSPPQLACGPTVHTCSENYVVKGLCELFNSNLRQPOK 120  
 80 GLQVPEAVNMSLGLSLATTSPPQLACGPTVHTCSENYVVKGLCELFNSNLRQPOK 137  
 121 FPLALRGCPQEDDIAFLDGGSIIPHFREMKSELVSTIMEQLKSKSTLFLSMQYSEEF 180  
 138 LPVSRQCEQEDIVFLDGGSIISRNFATMNFVRAVISQFORPSTQSLAQFSNKF 197  
 181 RHFTTKFQNNPNSLTKPTOLLGRTHFATGLKRVRELNFNTGARKNAKILFL 240  
 198 QHTFTFEEFRNTSNPLSLASVHQDGFYTTATQNVVHRLPHASYGARDAKILIVI 257  
 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDGAFSEKSRQELMTVASRPRDHFQAN 300  
 258 TDGKKGESLDYKDVIPMAADAAGIIRVAIGVGLAFONNSNKNELNDIAKSPQSHIFKE 317  
 301 NFEALXTVQNLREKIPALEGTOTGTSSEFHEMSQEGFAALTNGPLSTVGSYDNAG 360  
 318 DFDALNDIQNLREKIPALEGTOTGTSSEFHEMSQEGFAALTNGPLSTVGSYDNAG 377  
 361 GVFLYTSKEKSTFTINTRYDSMDNDAYLGYAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420

DB 378 GAFLYPPNMSPTFTNMNQENVDMDSDYLGYSTELALWKGVSQSLVGLAPRYOHIGKAVIFI 437  
 QY 421 QATGWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQSVCP 480  
 DB 438 QVSRQWRMKAQVIGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQSVCP 497  
 QY 481 PRGQARWQCDANLYGEGQDPWGFPAALTVLGDVNGDKLTDVAIGAPGEDNREGAVYLP 540  
 DB 498 PRGMR-RWCDVAVLYGEGQDPWGFPAALTVLGDVNGDKLTDVAIGAPGEDNREGAVYLP 556  
 QY 541 HGTSGSISPSHSORIASKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLLRSG 600  
 DB 557 HGVLPSPISPSHSORIASQLSSLOVFGQALSGQDLTQGLVDLAVGARGQVLLRTR 616  
 QY 601 PVLRYKATMEBNPREVAENVECNQVVKGEAGEVCLHVOKSTFDRREGQIOISVVT 660  
 DB 617 PVLWVGVSQFIPAEIPSAFECRQVVSQTLVQSNICLYIDRKNLGLSDLSQSVT 676  
 QY 661 YDLALDGRPHSRVAFNETNSTROTQVIGLGTOTCETLKLQLPNCHIEDVPSIVIRLNF 720  
 DB 677 LDALAPGRLSPRAI FOTTKRSLSRVRVLGKAKHCFENFLLPSCVEDSVIPIRLNF 736  
 QY 721 SLVGTPLGAFGNLRLPVLAEADQRLFTALPFEKNCNDNICODDLSITPSFMSLDCLVVG 780  
 DB 737 TLVGPPLAFNLRLPMLAALAAQRYFTASLPFEKNCADHICODNLSISFSPGLKSLVG 736  
 QY 781 GPRENVTVVRNCGDSYRTQVTFPFLDLSTKVKSTLQNSQSRQSWRLACESASSTEV 840  
 DB 797 SNLELNASVVMVNDGDSYGTITTFSPAGLSRYVAEGQKQQLSLHLTC--CSAPVG 854  
 QY 841 SCALXSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTWKTEF 900  
 DB 855 SQGTWSTSCRINHLIFRCGAQITFLATFDVSPRAVGLDRLLLIANVSSENNIPRTSKTIF 914  
 QY 901 QLELPVKAVYVTVSHCVSKYLNFTAS-ENTSRVMOHOYVSNLGRSLPLSLVLP 959  
 DB 915 QLELPVKAVYVTVSHCVSKYLNFTAS-ENTSRVMOHOYVSNLGRSLPLSLVLP 974  
 QY 960 VELNQTWDRPQVTFSENLSSTCHTKRPLSHSDFLAELKAPVNVNCISAVCORIQCDI 1019  
 DB 975 VELNQTWDRPQVTFSENLSSTCHTKRPLSHSDFLAELKAPVNVNCISAVCORIQCDI 1034  
 QY 1020 PFGIOEEFNATLKNLSFDWYIKTSHNHLIYSTABILNDVFTLLPCGQAFVRSQTE 1079  
 DB 1035 PSFSQDELDFTLKNLSFGVVRQILQKQSVSVSVABIIIFDTSVISOLPQGEAFMRAQTI 1094  
 QY 1080 TKVEPFYVNPPLIVGSSVGLLALLALITAAALYKLGFFKQYKQMMSE 1128  
 DB 1095 TVLEKYKVHNPILVIGSSIGGLLALLITAVLYKVGFFKQYKQMMSE 1143  
 RESULT 10  
 AAB07361  
 ID AAB07361 standard; protein; 1163 AA.  
 XX AC AAB07361;  
 XX DT 17-JAN-2001 (first entry)  
 XX DE Human CD11c protein sequence.  
 XX KW Human; macrophage infiltration inhibition; alpha\_d integrin;  
 KW leukocyte integrin; leu-CAM; leukocyte integrin; immune response;  
 KW inflammation; leukocyte adhesion deficiency; LAD; Type 1 diabetes;  
 KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;  
 KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;  
 KW rheumatoid arthritis; central nervous system injury; CD11c.  
 XX OS Homo sapiens.  
 XX WO200029446-A1.  
 XX PN





XX	Sun 2, Jay G;	Db	318	DFDALKDQOTQUREKI FPIEGTETSSSFELEMAQGFSAVFTDGPVLGAVGSFTWSG	377
PI	WPI; 2003-058520/05.	Qy	361	GVFLYTSKEKSTFFINMTREVDSDMNDAYLGAAAIILNRVQSLVGLGAPRYOHIGVAMPR	420
XX		Db	378	GAPLYFPNNSPTFINMSQENVMDRDSYLGSTELAKWGVQSLVGLGAPRYOHTGRAVFT	437
PT		Qy	421	QNTGHWESNANVKGTCIGAYFGASICSVDVDSNGSTDLVLICAPHYVYOTRGQGVSVCLP	480
XX		Db	438	QVSRWRKAEAVTGTQIGSYFGPSLCSVDVDSGSDTLVLIGPHHYEQTGCAQSVCLP	497
CC	The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially regulated in prostate cancer cells, which involves contacting a polypeptide differentially regulated in prostate cancer cells with a test agent under conditions effective for the test agent to modulate a biological activity of the polypeptide, and determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification of specific genes, and groups of genes, expressed in pathways physiologically relevant to prostate cancer permits the definition of functional and disease pathways and the delineation of targets in these pathways which are useful in diagnostic, therapeutic, and clinical applications. This is the amino acid sequence of a protein differentially regulated in prostate cancer	Qy	491	PRGQARWQCDVAVYBGCQWAGPFGAALTJVLGVNGDKLTADVAICAPCEENRGAYLIF	540
CC		Db	498	PRGNR-RWCDVAVYBGCQWAGPFGAALTJVLGVNGDKLTADVAICAPCEENRGAYLIF	556
CC		Qy	541	HCTSGSGLSPSHSQRACSKLSPLOYQCSLSCGODITDGLVDTLTGACGHVLLRSQ	600
CC		Db	557	HVLGFSIPSHSQRACSKLSPLOYQCSLSCGODITDGLVDTLTGACGHVLLRSQ	616
CC		Qy	601	PVLRYKALMEFNPREVAENVECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
CC		Db	617	PVLWVGSMQFIPABI PRSAFECREQVVEQTLVQSNICLVIDKSKNLLGSRDLQSSVT	676
CC		Qy	661	YDLALDSGRPHSRVAFENETKSTRQTOVLGTTCTETLKLQLPNCIEDPSPVILRLNF	720
CC		Db	677	LDLALDPGLSPRATFQETKRSLSRVRLVGLKAHCENFNLLPSCVDSVTPTILRLNF	736
CC		Qy	721	SLVGTPLSAFGNLPVLAEDAQRLFTALFPPEKQKCGNDNICQDDLSITFSFMSLCLVVG	780
CC		Db	737	TLVCKPLLAFLNLRPLAADAQRYFTASLPEKNCGADHICQDNLGIGSFPGLASLLVG	796
CC		Qy	781	GPREFVTVVRNDEGSYRQVTFPPFLDLSYRKVSTLQNRORSRWLACESASSTEV	840
CC		Db	797	SNLELNAEVMVMNDGEDSYGTTITFSHPAGLSYRYAEGQKQLRSLHLTCDSPAVG--	854
CC		Qy	841	SGALKSTSCSNHPIFFPENSVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF	900
CC		Db	855	SQGTWSTSCRINHIFRGAQITFLATFDVSPKAVLGDRLLLTANVSENNTPRTSKTTF	914
CC		Qy	901	QLELPKYAVVMYVTVSHGVSTKYLNFITAS-ENTSRVMOHQYQVNLGORSIPISLVFLVP	959
CC		Db	915	QLELPKYAVVMYVTVSHGVSTKYLNFITAS-ENTSRVMOHQYQVNLGORSIPISLVFLVP	974
CC		Qy	960	VLNQTVIWDPRPQVTFSENLSSTCTKRLSPSHSDFLAELKAPVNVNCSIAVCQRIQCDI	1019
CC		Db	975	VELNQEAVMVDVSVSLPONSLRCSSEKLAGSPASDFLAHQKNPVLDCSIAGCLFRCDV	1034
CC		Qy	1020	PFFGIQEEFNATLKNLSFDWYIKTSHNHLILIVTAEILFNDVSFTLLPFGQCAFVRSQTE	1079
CC		Db	1035	PSFSVQSELDFTLKNLSFGWVRLQKQKSVSVSVAITFTDTSVVSQLPQGEAFWRAQTT	1094
CC		Qy	1080	TKVPEFVNPDLPLIVGSSVCGHLLLLALITAAALKLGFKKQYKMMSE	1128
CC		Db	1095	TVLEKYKVNHTPLIVGSSIGCHLLLLALITAVLYKVGFFKQYKMMSE	1143
XX	RESULT 13				
XX	AA78166	ID	AA78166	standard; protein; 1161 AA.	
XX	AC	AA78166;			
XX	DT	28-DEC-1995	(first entry)		
XX	DE	Human beta-2 integrin alpha-d.			
XX	XX	Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis; inflammatory bowel disease; asthma.			
XX	OS	Homo sapiens.			
XX	PH	Key	Location/Qualifiers		
XX	FT	Domain	17..1108		

XX	Query Match	58.2%; Score 3421; DB 6; Length 1163;			
XX	Best Local Similarity	60.1%; Pred. No. 3.9e-274;			
XX	Matches	679; Conservative 141; Mismatches 303; Indels 6; Gaps 4;			
QY	1	FNLDTENMTFOBNARGFCQSVVLOGSRVYVVGAPQEIIVANQKSLYQCDYSTGSCBPI	60		
Db	20	FNLDTEELTAFRVDSAGFSDVVQYANSVWVGAPQKITAANQTCGLYQCYSTGACBPI	79		
QY	61	RLQVPEAVNMLGLSLAATSPQQLACGPTVHOTCSENTYVYKGLCFGLGSLNQPOK	120		
Db	80	GLQVPEAVNMLGLSLAATSPQQLACGPTVHOTCSENTYVYKGLCFGLGSLNQPOK	137		
QY	121	FPFALGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMBQKXKTLFSLMQYSEFP	180		
Db	138	LPVSRQECPRQEQDIVFLIDGSGSISSRNFAFMFRAVTSQFQRPSTQFSLMQFSNKP	197		
QY	181	RHETPEKQNNPNSLTKPTTOLLGRTHATGLRKVRELFINITNGAKNAKILILL	240		
Db	198	QTHLTPEERFRTNPLSLASVHLQDGYTTATQNVVHRLPHASYGARRDATKILLVI	257		
QY	241	TGCEKFGDPLGYEDVTELPEDRGVIRYVLGFGDAFSEKSRQELMTVASKPPRDRHVFOAN	300		
Db	258	TGCKEGDLDYKOVIPMDAAGIIRYALVGLAQNRNSWKEINDLASKPSQEHFKVE	317		
QY	301	NFEALKTQVQUREKIPALEGTQTGSSSFEHMSQEGFSAAITNGPLLLSTVGSDWAG	360		

FT	Region	/note= "extracellular domain"	
FT	150..352		
FT	/note= "this region is homologous to the insert common to Chla,b,c and may be a site for interaction with ICAM family proteins"		
FT	465..474		
FT	Binding-site	/note= "putative cation binding site"	
FT	518..527		
FT	Binding-site	/note= "putative cation binding site"	
FT	592..600		
FT	Binding-site	/note= "putative cation binding site"	
FT	1109..1128		
FT	Region	/note= "transmembrane region"	
FT	1129..1161		
FT	Domain	/note= "cytoplasmic domain"	
XX	NO9517412-A1.		
XX	29-JUN-1995.		
XX	21-DEC-1994;	94WO-US014832.	
XX	23-DEC-1993;	93US-00173497.	
XX	05-AUG-1994;	94US-00286889.	
XX	(ICOS-) ICOS CORP.		
XX	Gallatin WM, Van Der Vieren M;		
XX	WPI; 1995-240603/31.		
XX	N-PSDB; AAQ91712.		
XX	Alpha sub-unit polypeptide of human beta 2 integrin - used to identify potential anti-inflammatory agents, for the treatment of graft arteriosclerosis, inflammatory bowel disease, asthma, etc.		
XX	Claim 7; Page 82-87; 172pp; English.		
XX	A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha -TM1 was used to screen a human spleen cDNA library to identify clone 19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was manipulated to allow expression of recombinant alpha-d subunit in COS and CHO cells		
XX	Sequence 1161 AA;		
XX	Query Match	57.6%; Score 3388; DB 2; Length 1161;	
XX	Best Local Similarity	58.9%; Pred. No. 2.2e-271;	
XX	Matches	665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;	
Qy	1	FNLDENAMFQENARGGQSVVLOGSRVVGAPQEIIVAAQSGSLYQCDYSTGSCPI 60	
Db	17	FNLDVEERTIFQEDAGGQSVVQFGSRVLVVGAPLEVAANOTGRLYDCAATGMCQPI 76	
Qy	61	RLOVPEAVNMSGLSLAATSPQLAACGPTVHQTCSNTYVYKGLCFLFGSNLRQOPQK 120	
Db	77	PLHIRPEAVNMSGLTTLAASNGSRLACGPTLHRVCGENSYSKSGCLLGSRW-BIIQT 135	
Qy	121	PPALRCQEDSDIAPLIDGSGIIPHDPRMKELVSTIMEQLKSKTLPISIMQYSEEF 180	
Db	136	VPDAPTECPHQEMDIVFLIDGSGIDQNDENQMKFGVQAVMGQFEGDTLPALMQVSNLL 195	
Qy	181	RIHFTFKFQNNPNRSLIKPITOLLGRTHATGLRVRVRELFTNITNGARKNAFKILFL 240	
Db	196	KIHFTFTQFTSPSQSLVDPVQLKGLTFTATGILTVVQLFHHKNGARKSAKKILIVI 255	
Qy	241	TGDKFGDPIGYENVIPBELDEGVIRVILGFGAPRSEKROELNVTASPPDDHVPQAN 300	
Db	256	TGQKYDPLFYSVDPQAEAGIIRYAIAGVGHAFQGPQTARQELNITISSAPPDQHVFKVD 315	
Qy	301	NFEALKTVQQLREKI PAIEGTGTGSSSSEHEMSQEPGSAATISNGPILLSTVGSYDWAG 360	
Db	316	NFAALGSIQLOLQEKIYAVEGTQSRASSSFQHEMSQEGFTALTMDGLFLGAVGSFWSG 375	
Qy	361	GVFLYTSKEKSTPIINMTRVDSMDNDAYLGIAAAIILNPNVQSLVLCAPRYQHILGAMFR 420	
Db	376	GAFLYPPNPKSPFIINMSQENVDNRDSYLSYFELAKWGVQNLGAPRYQHTKAVIFT 435	
Qy	421	QNTGWESNANVKGTQIGAYFGASLCSDVDNSGSDTDLVILGAPHYVEQTRGGQSVCP 480	
Db	436	QVSRQWRKKAETVGTQIGSYFGASLCSDVDNSGSDTDLILGAPHYVEQTRGGQSVCP 495	
Qy	481	PRGQARWOCDAVLYGEOGPWGRFGAALTVDVNGDKLTVVAIGAPDEENRGAVILF 540	
Db	496	PRGQARWOCDAVLYGEOGPWGRFGAALTVDVNGDKLTVVAIGAPDEENRGAVILF 555	
Qy	541	HGTSGGISPSHSQRIAGSKLSPRLQYFGQSSGQDLMDCGLVDLTGAGQHVLLLRQ 600	
Db	556	HGASESGISPSHSQRIAGSKLSPRLQYFGQSSGQDLMDCGLVDLTGAGQHVLLLRQ 615	
Qy	601	PVLRYKAIMFNPVRVARNVFEQNDQVVKGEAGVRCVCLHVQKSTRDLRREGQTSVVT 660	
Db	616	PVLRYKAVAMRFSFVAVKAVRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673	
Qy	661	YDLALDSGRPHSRVFNENSTRTQTOVLGTQTCETLKLQLPNCIEBPVPIVLRNF 720	
Db	674	FDLALDPGLTSRALFNETKNPTLTRKTLGLGICETLKLPPCVDVDSPIILHNF 733	
Qy	721	SLVGTPLSAPGNLRLPVLAEQAORLFTALPPFPKNCQNDNICDDLSITFSFMSLDCLVVG 780	
Db	734	SLVREPIPSQNLRLPVLAVGSQDLFTASLPFEKNCQDGLCEGLGVTLISFGLQTLTVG 793	
Qy	781	GPREFNVTVTVNDGEDSVRTQVTFPEPLDLSYRVKVTI-ONORSQSRMLACESASTEV 840	
Db	794	SSLENLVITVWNAGDSYGVTVSYIYPAGLSHRVSGAQKQPHQSALRLACETV-PTED 852	
Qy	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANKVTENNMPRTNKEF 900	
Db	853	EG-LASSRCSVNHPIFHEGSGTFTVTFDVSYKATLGDRLMLRASASSENKASSKATF 911	
Qy	901	QLELPVKYAVNVVTSHGVSYKLYNF-TASENTSRVMOHOYQVNLGORSILPISLVLVP 959	
Db	912	QLELPVKYAVNVVTSHGVSYKLYNF-TASENTSRVMOHOYQVNLGORSILPISLVLVP 971	
Qy	960	VLNQTQVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISAVCORIQCDI 1019	
Db	972	VLLNGVAVVDVMEAPSQSL--PCVSEKRPQHSDFLTQISPSMLDCSIADCLQRCIDV 1029	
Qy	1020	PPFGIOREPNATLKNLSFDVYIKTSHNHLIVSTAEILPNDSTVTLPGQCAFVRSQTE 1079	
Db	1030	PSFSVQHELDFTLKGSLSPGWVRETQKKVLVSVSAEITFDTSVYSQLPQGEAFRAQME 1089	
Qy	1080	TKVPEPEVPNPLPLTVGSSVGLLALLALITLALYKLGFFKQYKDMMSR 1128	
Db	1090	MVLEDEVTNAPILINGSVGNALLALITLALYKLGFFKQYKDMMSR 1138	
XX	RESULT 14		
XX	AAW23049		
ID	AAW23049	standard; protein; 1161 AA.	
XX	AAW23049;		
XX	24-FEB-1998	(first entry)	
DE	Human beta 2 integrin alpha d subunit.		
KW	Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion; phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma; peoriasis; lung inflammation; acute respiratory distress syndrome; rheumatoid arthritis.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
PH	Domain	17..1108	

```

FT FT /label= Extracellular_domain
FT FT 150..352
FT FT /note= "region homologous to the I (insertion) domain
FT FT common to CD11a, CD11b and CD11c"
FT FT 1109..1128
FT FT /label= Transmembrane_domain
FT FT /note= "homologous to the human CD11c transmembrane
FT FT region"
FT FT 1129..1161
FT FT /label= Cytoplasmic_domain
XX XX
XX XX WO9731039-A1.
XX XX
XX XX 28-AUG-1997.
XX XX
XX XX 24-FEB-1997; 97WO-US002713.
XX XX
XX XX 22-FEB-1996; 96US-00605672.
XX XX
XX XX (ICOS-) ICOS CORP.
XX XX
XX XX Gallatin WM, Van Der Vieren M;
XX XX
XX XX WPI; 1997-435154/40.
XX XX
XX XX DR N-PSDB; AAT79220.
XX XX
XX XX Hybridoma 199M and antibody secreted by it - specific for new rat beta2
XX XX integrin subunit, useful to detect subunit in cells and modulate its
XX XX activity.
XX XX
XX XX Example 5; Page 116-120; 222pp; English.
XX XX
XX XX This polypeptide comprises a novel human beta 2 integrin subunit,
XX XX designated alpha d. Its sequence was deduced from a cDNA clone (see
XX XX AAT73220) isolated from a spleen cDNA library. Alpha d is involved in
XX XX cell migration, phagocytosis and cell-cell interaction. Recombinant alpha
XX XX d polypeptides can be expressed in transformed host cells for use in
XX XX assays for identifying antibodies or other compounds that modulate alpha
XX XX d activity or which modulate the interaction between alpha d and a
XX XX ligand, for treating or preventing diseases in which macrophages are
XX XX implicated. Treatment is applicable to disease states in which alpha d
XX XX binding, or localised accumulation of cells which express alpha d, is
XX XX implicated such as type I diabetes, atherosclerosis, multiple
XX XX sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
XX XX distress syndrome and rheumatoid arthritis
XX XX
XX XX Sequence 1161 AA;
XX XX
XX XX Query Match 57.6%; Score 3388; DB 2; Length 1161;
XX XX Best Local Similarity 58.9%; Pred. No. 2.2e-271;
XX XX Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6
XX XX
XX QY 1 FNLDYENAMTFQENARGFQSQVVLQGSRRVVGAPQBIVAANQSGLSYQCDYSTGSCPEI 60
XX DB 17 FNLDVEEPTIFQDAGGFGQSVVQGSRLVVGAPLEVAANQTGRLYDCAATGMCQPI 76
XX QY 61 RLQVPVEAVNMSLGLSLAATSPPOLLAGTGVHTQCSNTYVKGICFLPGSNLRQPOK 120
XX DB 77 PPHIRPEAVNMSLGLTLAASNGSRELLACGPTLHRVCGENSYSKGSCLLGSRW-ETIOT 135
XX QY 121 FPEALRGCPQEDSDIAFLIDGSGSTIIPHDPRMKELVSTIMEOLKSKTLFSLMQYSSEF 180
XX DB 136 VQDATPECPHQENDIVFLIDGSGSIDQNDFNQMGFTQAQVWQPEGTDTFALMQYSNLL 195
XX QY 181 RHFTTKPQNPNPRSLIKPIITOLLGTHTATGLRKVKVRELNIINGARONAFKILFL 240
XX DB 196 KIHFTPTQRTSPSQSLVDPTVOLKGLTFTATGILTVVTLFHHNKGARKSAKKILIVI 255
XX QY 241 TIGKFGDPLGVEDYVPELDREGVIRYVLGPDGAFRSEKSRQELNTVASKPRDPAVFQAN 300
XX DB 256 TDGCKYKPLEYSDVIPQEKAGHIRVAIGVHFAFQGPPTARQELNTISSAPQDHFVKVD 315
XX QY 301 NFEALXTQNLQREKIFAEGTQTGCGSSSSPEHMSQGFSAAITSGNPLLTSTGVSYDNAG 360

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:07:20 ; Search time 18.3484 Seconds  
(without alignments)  
3199.127 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGAEFQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/2/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/prodata/2/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5839	99.3	1153	1	US-08-173-497-3
2	5839	99.3	1153	1	US-08-286-889-3
3	5839	99.3	1153	1	US-08-485-618-3
4	5839	99.3	1153	1	US-08-362-652-3
5	5839	99.3	1153	2	US-08-605-672-3
6	5839	99.3	1153	2	US-08-482-293A-3
7	5839	99.3	1153	2	US-08-943-363-3
8	5839	99.3	1153	3	US-09-193-043-3
9	5839	99.3	1153	4	US-09-688-307A-3
10	5839	99.3	1153	4	US-09-350-259-3
11	5808.5	98.8	1152	2	US-08-476-062A-43
12	5808.5	98.8	1152	5	PCT-US96-01314-43
13	5808.5	98.8	1152	2	542339-2
14	3446	58.6	1163	2	US-08-476-062A-44
15	3446	58.6	1163	5	PCT-US96-01314-44
16	3423	58.2	1163	1	US-08-173-497-4
17	3423	58.2	1163	1	US-08-286-889-4
18	3423	58.2	1163	1	US-08-485-618-4
19	3423	58.2	1163	1	US-08-362-652-4
20	3423	58.2	1163	2	US-08-605-672-4
21	3423	58.2	1163	2	US-08-482-293A-4
22	3423	58.2	1163	2	US-08-943-363-4
23	3423	58.2	1163	3	US-09-193-043-4
24	3423	58.2	1163	4	US-09-688-307A-4
25	3423	58.2	1163	4	US-09-350-259-4
26	3388	57.6	1161	1	US-08-173-497-2
27	3388	57.6	1161	1	US-08-286-889-2

28 3388 57.6 1161 1 US-08-485-618-2 Sequence 2, Appli  
29 3388 57.6 1161 1 US-08-362-652-2 Sequence 2, Appli  
30 3388 57.6 1161 2 US-08-605-672-2 Sequence 2, Appli  
31 3388 57.6 1161 2 US-08-482-293A-2 Sequence 2, Appli  
32 3388 57.6 1161 2 US-08-943-363-2 Sequence 2, Appli  
33 3388 57.6 1161 3 US-09-193-043-2 Sequence 2, Appli  
34 3388 57.6 1161 4 US-09-688-307A-2 Sequence 2, Appli  
35 3388 57.6 1161 1 US-08-485-618-99 Sequence 99, Appli  
36 3372.5 57.4 1161 2 US-08-605-672-99 Sequence 99, Appli  
37 3372.5 57.4 1161 2 US-08-482-293A-99 Sequence 99, Appli  
38 3372.5 57.4 1161 2 US-08-943-363-99 Sequence 99, Appli  
39 3372.5 57.4 1161 3 US-09-193-043-99 Sequence 99, Appli  
40 3372.5 57.4 1161 4 US-09-688-307A-99 Sequence 99, Appli  
41 3372.5 57.4 1161 3 US-09-350-259-99 Sequence 55, Appli  
42 3372.5 57.4 1161 4 US-09-193-043-55 Sequence 55, Appli  
43 3213.5 54.7 1161 4 US-09-688-307A-55 Sequence 55, Appli  
44 3213.5 54.7 1161 4 US-09-350-259-55 Sequence 55, Appli  
45 3213.5 54.7 1161 4 US-09-688-307A-55 Sequence 55, Appli

#### ALIGNMENTS

RESULT 1  
US-08-173-497-3  
; Sequence 3, Application US/08173497  
; Patent No. 5437958  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha  
; TITLE OF INVENTION: Subunit  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 S. Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/173,497  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5437958and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 278666/31363  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-173-497-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANORGLYQCDYSTGSCBPI 60  
|||||  
DB 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANORGLYQCDYSTGSCBPI 76

```
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTILFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTILFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPTITQLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNRSLIKPTITQLGRTHATGLRKVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYVIGDFAFSEKSRQBLNTVASPPRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRVYVIGDFAFSEKSRQBLNTIASKPPRDHVFOVN 316
QY 301 NFALKTVQNLREKIPALRGTOGTGSSSPHEHMSQEGFSAATISNGPLLSSTVGSYDWAG 360
Db 317 NFALKTVQNLREKIPALRGTOGTGSSSPHEHMSQEGFSAATISNGPLLSSTVGSYDWAG 376
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDVGLGYAAAILLRNRVQSLVGLAPRYOHIGLVAMER 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDVGLGYAAAILLRNRVQSLVGLAPRYOHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCPIL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCPIL 496
QY 481 PRGORAWOCDVLYGEOGPKWGFAGALTIVLGVNGDKLTDVAIGARBEEDNRGAVILF 540
Db 497 PRGORAWOCDVLYGEOGPKWGFAGALTIVLGVNGDKLTDVAIGARBEEDNRGAVILF 556
QY 541 HGTSGSGISPSHSORIASGSLPRLOVFGQSLGGQDLTMDGLVDTLVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGSLPRLOVFGQSLGGQDLTMDGLVDTLVGAQGHVLLRSQ 616
QY 601 PVLRLVXALMEFNRVARNVFCNDQVYVKGKEAGRVVCLHVOKSTRDLREBQIQSVVT 660
Db 617 PVLRLVXALMEFNRVARNVFCNDQVYVKGKEAGRVVCLHVOKSTRDLREBQIQSVVT 676
QY 661 YDLALDSGRPHSAVFNEMTKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSAVFNEMTKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAPGNLRPVLAEDAQLFTALPPFEXNCGNDNICQDDLSITFSPMSLDCLVVG 780
Db 737 SLVGTPLSAPGNLRPVLAEDAQLFTALPPFEXNCGNDNICQDDLSITFSPMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLACESASSTEV 856
QY 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASIGKNLLKANTYSENMPRTNKTFF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASIGKNLLKANTYSENMPRTNKTFF 916
QY 901 QLELPVKYAVVMVTSHGUSTKYLNPTASNTSRVMQHOYVSNLQORSIPISLVLPLVPV 960
Db 917 QLELPVKYAVVMVTSHGUSTKYLNPTASNTSRVMQHOYVSNLQORSIPISLVLPLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1036
QY 1021 FFGIQEFNATLKNLSFDWYIKTSNNHLLIVSTASILNDSVFTLLPCQGAFVRSQTF 1080
Db 1037 FFGIQEFNATLKNLSFDWYIKTSNNHLLIVSTASILNDSVFTLLPCQGAFVRSQTF 1096
QY 1081 KVEPFFVPPLPLIVGSSVGGILLLALITAAALYKLGFFRQYKDMWSEGGPPCAEPQ 1137
Db 1097 KVEPFFVPPLPLIVGSSVGGILLLALITAAALYKLGFFRQYKDMWSEGGPPCAEPQ 1153
```

## RESULT 2

```
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3
```

Query Match 99.3%; Score 5839; DB 1; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 ENLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQBIVAANORGSLYCCDYSTGSCPEI 60
Db 17 ENLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQBIVAANORGSLYCCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTILFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTILFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPTITQLGRTHATGLRKVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNRSLIKPTITQLGRTHATGLRKVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVYVIGDFAFSEKSRQBLNTVASKPPRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRVYVIGDFAFSEKSRQBLNTIASKPPRDHVFOVN 316
QY 301 NFALKTVQNLREKIPALRGTOGTGSSSPHEHMSQEGFSAATISNGPLLSSTVGSYDWAG 360
Db 317 NFALKTVQNLREKIPALRGTOGTGSSSPHEHMSQEGFSAATISNGPLLSSTVGSYDWAG 376
```

QY 361 GVPLVTSKEKSTFNNTRVDSMDNDAYLGAAAIILRNVRQSLVIGAPRYQHIGLVAMER 420  
Db 377 GVPLVTSKEKSTFNNTRVDSMDNDAYLGAAAIILRNVRQSLVIGAPRYQHIGLVAMER 436  
QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVQVPL 480  
Db 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVQVPL 496  
QY 481 PRGORARWQCDVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
Db 497 PRGORARWQCDVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556  
QY 541 HTSGSGISPSHSQRIAGSKSPRIQYFGQSLGSGQDLTMDGLVDTLVGAQGHVLLRSQ 600  
Db 557 HTSGSGISPSHSQRIAGSKSPRIQYFGQSLGSGQDLTMDGLVDTLVGAQGHVLLRSQ 616  
QY 601 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 617 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETNSTRTQTQVLGTQTCETILKQLPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRVAFNETNSTRTQTQVLGTQTCETILKQLPNCIEDPVSPIVLRNF 736  
QY 721 SLVGTPLSAGNLRPVLAEDAQRLLFTALPPFEKNCNDNI CODDLSITFSFMSLDCLVYG 780  
Db 737 SLVGTPLSAGNLRPVLAEDAQRLLFTALPPFEKNCNDNI CODDLSITFSFMSLDCLVYG 796  
QY 781 GRPFRNVTVVNDGEDSVRTQVTFEPLDLSYRKVSTLQNRORSWRLACESASSTEV 840  
Db 797 GRPFRNVTVVNDGEDSVRTQVTFEPLDLSYRKVSTLQNRORSWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRINTKEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRINTKEF 916  
QY 901 QLELPVKAVMVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLPISVFLVPV 960  
Db 917 QLELPVKAVMVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLPISVFLVPV 976  
QY 961 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPDSHDFLAEIRKAPVVCNSIAYCQRIQCDIP 1020  
Db 977 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPDSHDFLAEIRKAPVVCNSIAYCQRIQCDIP 1036  
QY 1021 PFGIOEFNATLKGNSLFDWIKTSHNHLIVSTHAILFNDSVFTLLPGQAFVRSQTEF 1080  
Db 1037 PFGIOEFNATLKGNSLFDWIKTSHNHLIVSTHAILFNDSVFTLLPGQAFVRSQTEF 1096  
QY 1081 KYEPEVENPPLIVGSSVGGLLALITALYKLGPFKROYKDMMSGGPPGABPO 1137  
Db 1097 KYEPEVENPPLIVGSSVGGLLALITALYKLGPFKROYKDMMSGGPPGABPO 1153

RESULT 3  
US-08-485-618-3  
; Sequence 3, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5728533el Humar 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485.618  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32797  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-618-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGCGQSVVQLQGRVVVVGAPQBIIVAAORGSIVQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGCGQSVVQLQGRVVVVGAPQBIIVAAORGSIVQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKELCFLFGSNLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKELCFLFGSNLRQOPQK 136  
QY 121 PPEALRCQPEDSDIAPLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 180  
Db 137 PPEALRCQPEDSDIAPLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 196  
QY 181 RIHPTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240  
Db 197 RIHPTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILV 256  
QY 241 TDGEKFGDPLGYEDVTPELDREGVIRVYLGFGDAPRSEKSRQBELNTVASPPDRHVPQAN 300  
Db 257 TDGEKFGDPLGYEDVTPELDREGVIRVYLGFGDAPRSEKSRQBELNTIASKPPDRHVPQV 316  
QY 301 NFEALKTQNLREKIPALRGTTGSSSSPEHEMSQEGPSAALTNSGPLSTVGSYDWAG 360  
Db 317 NFEALKTQNLREKIPALRGTTGSSSSPEHEMSQEGPSAALTNSGPLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILRNVRQSLVIGAPRYQHIGLVAMER 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILRNVRQSLVIGAPRYQHIGLVAMER 436  
QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVQVPL 480  
Db 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVQVPL 496  
QY 481 PRGORARWQCDVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
Db 497 PRGORARWQCDVLYGEGQGPNGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556  
QY 541 HTSGSGISPSHSQRIAGSKSPRIQYFGQSLGSGQDLTMDGLVDTLVGAQGHVLLRSQ 600

557 HGTSGGISHSHQRTAGSKLSPLOYFGQSLSGGDLTWDGLVDTVGAQHVLRLRQ 616  
601 PVLVRKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVQKSTRDLRBCQIQSVVT 660  
617 PVLVRKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVQKSTRDLRBCQIQSVVT 676  
661 YDLALDSGRPHSRVAVNETKNSRRTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNP 720  
677 YDLALDSGRPHSRVAVNETKNSRRTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNP 736  
721 SLVGTPLSAFNGNLRPVLAEDAQLFTALPPFERKNCNDNICODDLSITFSFMSLDCLVVG 780  
737 SLVGTPLSAFNGNLRPVLAEDAQLFTALPPFERKNCNDNICODDLSITFSFMSLDCLVVG 796  
781 GPREENVTVVRNDGSDSYRTQVTPPEPDLISYKYSTLQNSORSWRLACESASTEV 840  
797 GPREENVTVVRNDGSDSYRTQVTPPEPDLISYKYSTLQNSORSWRLACESASTEV 856  
841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 900  
857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 916  
901 QLELPKYAYVMVTSHGVSUKYINFTASNTSRVQHQQVQVSNLQORSLPISLVLVYV 960  
917 QLELPKYAYVMVTSHGVSUKYINFTASNTSRVQHQQVQVSNLQORSLPISLVLVYV 976  
961 RLQNTWIDRPOVTFSENSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1020  
977 RLQNTWIDRPOVTFSENSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1036  
1021 FPGIOBEFNATLGNLSDFYIYKTSNNHLLIVSTABILFNDVFTLLPQGGAFVRSQTET 1080  
1037 FPGIOBEFNATLGNLSDFYIYKTSNNHLLIVSTABILFNDVFTLLPQGGAFVRSQTET 1096  
1081 KVEPFEVNPPLPIVSSVGLLLALITAAALYKLGFKRQYKDMSEGGPPGABPQ 1137  
1097 KVEPFEVNPPLPIVSSVGLLLALITAAALYKLGFKRQYKDMSEGGPPGABPQ 1153

RESULT 4  
US-08-362-652-3  
Sequence 3, Application US/08362652  
Patent No. 5766850  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 5300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/362,652  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32391  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-362-652-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSVVQLQGSRVVVGAPQEIIVAAANORGLSYCDYSTGSCSPI 60  
DB 17 FNLDTENAMTQENARGFGQSVVQLQGSRVVVGAPQEIIVAAANORGLSYCDYSTGSCSPI 76  
QY 61 RLOVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKLCFLPGSNLRQOQOK 120  
DB 77 RLOVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKLCFLPGSNLRQOQOK 136  
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMOYSEEP 180  
DB 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMOYSEEP 196  
QY 181 RIHFTFKFQNNPNRSLIKDITQLLGHSTHTATGKRVKRVRELFNITNGARKNAKFLVLL 240  
DB 197 RIHFTFKFQNNPNRSLIKDITQLLGHSTHTATGKRVKRVRELFNITNGARKNAKFLVLL 256  
QY 241 TDGEKFGPLGVEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPRDRHVFOAN 300  
DB 257 TDGEKFGPLGVEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPRDRHVFOAN 316  
QY 301 NFEALKTQNLREKIPAEIGTQSSSSPEHEMSQEGFSAITNSGILLSTVGSYDWAG 360  
DB 317 NFEALKTQNLREKIPAEIGTQSSSSPEHEMSQEGFSAITNSGILLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILNRNRVOSLVLCAPRYOHIGLVAMPR 420  
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILNRNRVOSLVLCAPRYOHIGLVAMPR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYVQTRGGQVSVCP 480  
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLICAPHYVQTRGGQVSVCP 496  
QY 481 PRGQARWQCDVAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVALGAPGEEDNRGAVYLF 540  
DB 497 PRGQARWQCDVAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVALGAPGEEDNRGAVYLF 556  
QY 541 HGTSGGISHSHQRTAGSKLSPLOYFGQSLSGGDLTWDGLVDTVGAQHVLRLRQ 600  
DB 557 HGTSGGISHSHQRTAGSKLSPLOYFGQSLSGGDLTWDGLVDTVGAQHVLRLRQ 616  
QY 601 PVLVRKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVQKSTRDLRBCQIQSVVT 660  
DB 617 PVLVRKAIEMFNPREVARNVFCNDQVYKGEAGVRVCLHVQKSTRDLRBCQIQSVVT 676  
QY 661 YDLALDSGRPHSRVAVNETKNSRRTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNP 720  
DB 677 YDLALDSGRPHSRVAVNETKNSRRTQVGLTQTCETLKLQLPNCIBDPVSPVILRLNP 736  
QY 721 SLVGTPLSAFNGNLRPVLAEDAQLFTALPPFERKNCNDNICODDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFNGNLRPVLAEDAQLFTALPPFERKNCNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GPREENVTVVRNDGSDSYRTQVTPPEPDLISYKYSTLQNSORSWRLACESASTEV 840

Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASIGNKLLKANKVTSNNMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASIGNKLLKANKVTSNNMPRTNKTEF 916  
Qy 901 QLELPVKYAVYMWVTSKHGVSSTKYNLTASNTSRVMQHOVQVSNLQORSIPISLVLFPV 960  
Db 917 QLELPVKYAVYMWVTSKHGVSSTKYNLTASNTSRVMQHOVQVSNLQORSIPISLVLFPV 976  
Qy 961 RLNQTVINDRPOVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1020  
Db 977 RLNQTVINDRPOVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1036  
Qy 1021 FPGIOSEFNATLKNLSFOWYKTSNHNLLIVSTAEILFNDVSPTLPGOGAFVRSOTET 1080  
Db 1037 FPGIOSEFNATLKNLSFOWYKTSNHNLLIVSTAEILFNDVSPTLPGOGAFVRSOTET 1096  
Qy 1081 KVEPFVEPNPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMKSEGGPPGABEQ 1137  
Db 1097 KVEPFVEPNPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMKSEGGPPGABEQ 1153

RESULT 5

US-08-605-672-3  
; Sequence 3, Application US/08605672  
; Patent No. 5817515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402

Query Match 99.3%; Score 5839; DB 2; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQORSLVOCYSTGSCBPI 60  
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQORSLVOCYSTGSCBPI 76  
Qy 61 RLOVPEAVNNLSGLSLAANTSPPOLLAGCPVHOTCSNTVYKGLCFGLFNSLQOQPOK 120  
Db 77 RLOVPEAVNNLSGLSLAANTSPPOLLAGCPVHOTCSNTVYKGLCFGLFNSLQOQPOK 136  
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPDPRMKELVSTIMEQLKSKTLPMSLMOYSBEF 180  
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPDPRMKELVSTIMEQLKSKTLPMSLMOYSBEF 196  
Qy 181 RIHPTFKFQONPNRSLIKBITOLGRTHATGKRVVRELFTNITNGARKNAFILL 240  
Db 197 RIHPTFKFQONPNRSLIKBITOLGRTHATGKRVVRELFTNITNGARKNAFILL 256  
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELMTASKPRDHVQFN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELMTASKPRDHVQFN 316  
Qy 301 NFEALKTQONQUREKIPAIETGTCSSSSFEHMSQBSAAITNGPILLSVGSVDWAG 360  
Db 317 NFEALKTQONQUREKIPAIETGTCSSSSFEHMSQBSAAITNGPILLSVGSVDWAG 376  
Qy 361 GYPLVTSKEKSTFINMTKRVDSMDNDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420  
Db 377 GYPLVTSKEKSTFINMTKRVDSMDNDAYLGAAAAIILNRVQSLVLCAPRYQHIGLVAMFR 436  
Qy 421 QNTGMBESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLICAGPHYBETRGQSVVCP 480  
Db 437 QNTGMBESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLICAGPHYBETRGQSVVCP 496  
Qy 481 PRGQARWQCDVAVLXGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAIVLF 540  
Db 497 PRGQARWQCDVAVLXGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAIVLF 556  
Qy 541 HGTSGSISPSHSQRIAGSKISPRIOYFGQSLSGGQDLTMDGLVLDITVGAQHVLLLRQ 600  
Db 557 HGTSGSISPSHSQRIAGSKISPRIOYFGQSLSGGQDLTMDGLVLDITVGAQHVLLLRQ 616  
Qy 601 PVLRVKATMBENPREVARNVPECNDQVVKGEAGVRVCLHVQKSTRDLRREGQIQSVVT 660  
Db 617 PVLRVKATMBENPREVARNVPECNDQVVKGEAGVRVCLHVQKSTRDLRREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRAVFNENKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720  
Db 677 YDLALDSGRPHSRAVFNENKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 736  
Qy 721 SLVGTPLSAFGNLRPVLAEDAORLFTALFPPEKNGNDNICODDLSITFSEMSLQCLVVG 780  
Db 737 SLVGTPLSAFGNLRPVLAEDAORLFTALFPPEKNGNDNICODDLSITFSEMSLQCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASIGNKLLKANKVTSNNMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASIGNKLLKANKVTSNNMPRTNKTEF 916  
Qy 901 QLELPVKYAVYMWVTSKHGVSSTKYNLTASNTSRVMQHOVQVSNLQORSIPISLVLFPV 960  
Db 917 QLELPVKYAVYMWVTSKHGVSSTKYNLTASNTSRVMQHOVQVSNLQORSIPISLVLFPV 976  
Qy 961 RLNQTVINDRPOVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1020  
Db 977 RLNQTVINDRPOVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1036

QY 1021 PFGIOEFNATLKNLSFDMVYKTSNHLIIVSTABILLFNDSTVFTLLPGQAFVRSQTET 1080  
DB 1037 PFGIOEFNATLKNLSFDMVYKTSNHLIIVSTABILLFNDSTVFTLLPGQAFVRSQTET 1096  
QY 1081 KVBEPFVNPPLIVGSSVGLLALITAAALYKLGFFKQYKQKDMSEGGPPGAEPO 1137  
DB 1097 KVBEPFVNPPLIVGSSVGLLALITAAALYKLGFFKQYKQKDMSEGGPPGAEPO 1153

## RESULT 6

US-08-482-293A-3  
; Sequence 3, Application US/08482293A  
; Patent No. 5831029  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,293A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 278666/32584  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-482-293A-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARGFQSVVQLQGGRRVVVGAPQEIIVAAQORGLYQCDYSTGSCPEI 60  
DB 17 ENLDTENAMTFQENARGFQSVVQLQGGRRVVVGAPQEIIVAAQORGLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGKLCFLFGSNLRQQPQK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGKLCFLFGSNLRQQPQK 136  
QY 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKFEVSTVMEQLKSKTFLSLMQYSEEP 180

## RESULT 7

US-08-943-363-3  
; Sequence 3, Application US/08943363  
; Patent No. 5837478  
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: US/08/943,363  
FILING DATE: 5-AUG-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-943-363-3

Query Match  
Best local Similarity 99.3%; Score 5839; DB 2; Length 1153;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCEPI 60  
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCEPI 76

QY 61 RLQVPVAVNMSGLSLAATTSPPQLACGPTVHQCSTENTYKGLCFPLGSLNRQOPK 120  
DB 77 RLQVPVAVNMSGLSLAATTSPPQLACGPTVHQCSTENTYKGLCFPLGSLNRQOPK 136

QY 121 FPEALRCGPEDSDIAFLDGSIIIPDFFRMKELVSTIMEOLKSKTFLPSLMQYSEEP 180  
DB 137 FPEALRCGPEDSDIAFLDGSIIIPDFFRMKELVSTIMEOLKSKTFLPSLMQYSEEP 196

QY 181 RIHPTFKEPQNNPRLIPIKPIITOLLGRTATGLRKVRVRELFININGARKNAFKILFL 240  
DB 197 RIHPTFKEPQNNPRLIPIKPIITOLLGRTATGLRKVRVRELFININGARKNAFKILVI 256

QY 241 TDCBKFGDPLGYEDVIEPDLREGVIRVVLGFGAPFSEKSRQELNTVASPPRDHVFQAN 300  
DB 257 TDGKFGDPLGYEDVIEPDLREGVIRVVLGFGAPFSEKSRQELNTVASPPRDHVFQAN 316

QY 301 NFEALKTVQNLREKIFAIECTQTGSSSSSEHEMSQEGFSAATSNGLSTVGSYDWAG 360  
DB 317 NFEALKTVQNLREKIFAIECTQTGSSSSSEHEMSQEGFSAATSNGLSTVGSYDWAG 376

RESULT 8  
US-09-193-043-3  
; Sequence 3, Application US/09193043  
; Patent No. 6251395  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6251395e1 Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER FILING DATE: 1997-10-03

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLAGAPRYOHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLAGAPRYOHIGLVAMFR 436

QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNASTDLVLIGAPHYTYQTGGQSVQCP 480  
DB 437 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNASTDLVLIGAPHYTYQTGGQSVQCP 496

QY 481 PRGQARWQCDVLYGEGQQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGQARWQCDVLYGEGQQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGCHVLLASQ 600  
DB 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTVGAGCHVLLASQ 616

QY 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 660  
DB 617 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 676

QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQENCIEDPSPVILRLNF 720  
DB 677 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQENCIEDPSPVILRLNF 736

QY 721 SLVGTPLSAPGNLRLPVLAEADQRLFTALFPPEKKGNDNICODDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAPGNLRLPVLAEADQRLFTALFPPEKKGNDNICODDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTTVNRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
DB 797 GPREFNVTTVNRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900  
DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916

QY 901 QLELFPKYAVVMVTVSHGVSTKYLNTASNTSRVMOHQYQVSNIGQSLPISLFLVFPV 960  
DB 917 QLELFPKYAVVMVTVSHGVSTKYLNTASNTSRVMOHQYQVSNIGQSLPISLFLVFPV 976

QY 961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHDSFLAELRKAPVNCSTAVCQRIQCDDIP 1020  
DB 977 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHDSFLAELRKAPVNCSTAVCQRIQCDDIP 1036

QY 1021 FPGIOEFNATLKGNLSPDWIKTSHNHLIIVSTAEIIFNDSVFTLLPQCGAFVRSQSTET 1080  
DB 1037 FPGIOEFNATLKGNLSPDWIKTSHNHLIIVSTAEIIFNDSVFTLLPQCGAFVRSQSTET 1096

QY 1081 KVEPFEVNPPLIIVGSSVGGLLILALITAAALYKLGPPKQYKDMMSGGPPGABPQ 1137  
DB 1097 KVEPFEVNPPLIIVGSSVGGLLILALITAAALYKLGPPKQYKDMMSGGPPGABPQ 1153



Db 317 NFEALKTIONQREKIFAIEGTQGTSSSFEHMSQGFSAAITNSGPLLSTVGSYDAG 376  
Qy 361 GYFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYOHIGLVAMER 420  
Db 377 GYFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYOHIGLVAMER 436  
Qy 421 QNTGWMESNANVKGTOIGAYFGASLCSDVDNGSTDLVLI GAPHYETRGQGVSVCP 480  
Db 437 QNTGWMESNANVKGTOIGAYFGASLCSDVDNGSTDLVLI GAPHYETRGQGVSVCP 496  
Qy 481 PRGQARWQCDVAVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWQCDVAVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
Qy 541 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 616  
Qy 601 PVLRVKAIKMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRRCQIOSVVT 660  
Db 617 PVLRVKAIKMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRRCQIOSVVT 676  
Qy 661 YDLALDSGRPHSRVAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRVAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736  
Qy 721 SLVGTPLSAFNGLRPVLAEDAORLFTALPPFEKNCNDNI CQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNGLRPVLAEDAORLFTALPPFEKNCNDNI CQDDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPFPENSEVTFNITFDVDSKASLGKLLKANKYTSNNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPFPENSEVTFNITFDVDSKASLGKLLKANKYTSNNMPTNKTEF 916  
Qy 901 QLELPVKYAVYVWVTSHGVSSTKYINFTASNTSRVMOCHOVQVSNLQORSLSPLVFLVPV 960  
Db 917 QLELPVKYAVYVWVTSHGVSSTKYINFTASNTSRVMOCHOVQVSNLQORSLSPLVFLVPV 976  
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCIP 1036  
Qy 1021 PFGIOEFNATLKNLSFDVYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTET 1080  
Db 1037 PFGIOEFNATLKNLSFDVYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTET 1096  
Qy 1081 KVEPPEFVNPPLTVGSSVGGMLLALITAAALYKLGFFKQYKDMMSBEGPPGAEFQ 1137  
Db 1097 KVEPPEFVNPPLTVGSSVGGMLLALITAAALYKLGFFKQYKDMMSBEGPPGAEFQ 1153

## RESULT 10

US-09-350-259-3  
; Sequence 3, Application US/09350259  
; Patent No. 6620915  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; TITLE OF INVENTION: No. 6620915el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 4; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLVQCYSTSCBPI 60  
Db 17 FNLDTENAMTQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLVQCYSTSCBPI 76  
Qy 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTVVKGLCFLPGSNLRFQOQK 120  
Db 77 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTVVKGLCFLPGSNLRFQOQK 136  
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLPFSLMOYSBEP 180  
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLPFSLMOYSBEP 196  
Qy 181 RIHFTFKFQNNPNSRLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAKILPVL 240  
Db 197 RIHFTFKFQNNPNSRLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAKILPVL 256  
Qy 241 TDGEFGFGLGYEDVPIBLDREGVIRYVGLFGDAFSEKSRQELNVTASKPPOHVFQAN 300  
Db 257 TDGEFGFGLGYEDVPIBLDREGVIRYVGLFGDAFSEKSRQELNVTASKPPOHVFQAN 316  
Qy 301 NFEALKTIONQREKIFAIEGTQGTSSSFEHMSQGFSAAITNSGPLLSTVGSYDAG 360  
Db 317 NFEALKTIONQREKIFAIEGTQGTSSSFEHMSQGFSAAITNSGPLLSTVGSYDAG 376  
Qy 361 GYFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYOHIGLVAMER 420  
Db 377 GYFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYOHIGLVAMER 436  
Qy 421 QNTGWMESNANVKGTOIGAYFGASLCSDVDNGSTDLVLI GAPHYETRGQGVSVCP 480  
Db 437 QNTGWMESNANVKGTOIGAYFGASLCSDVDNGSTDLVLI GAPHYETRGQGVSVCP 496  
Qy 481 PRGQARWQCDVAVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWQCDVAVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
Qy 541 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 616  
Qy 601 PVLRVKAIKMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRRCQIOSVVT 660  
Db 617 PVLRVKAIKMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRRCQIOSVVT 676  
Qy 661 YDLALDSGRPHSRVAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRVAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736  
Qy 721 SLVGTPLSAFNGLRPVLAEDAORLFTALPPFEKNCNDNI CQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNGLRPVLAEDAORLFTALPPFEKNCNDNI CQDDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPFPENSEVTFNITFDVDSKASLGKLLKANKYTSNNMPTNKTEF 900

Db 857 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEP 916  
Qy 901 QLELPVKYAVYVWVTSHGVSCTKYLNTASNTSRVMQHOYOVNSLQORSIPISLVELVPV 960  
Db 917 QLELPVKYAVYVWVTSHGVSCTKYLNTASNTSRVMQHOYOVNSLQORSIPISLVELVPV 976  
Qy 961 RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQCDIP 1020  
Db 977 RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQCDIP 1036  
Qy 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEP 1080  
Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEP 1096  
Qy 1081 KVEPFEVNPFLPLIVGSSVGLLALLITAALYKLGFFKQYKDMMSGEGPPGABPQ 1137  
Db 1097 KVEPFEVNPFLPLIVGSSVGLLALLITAALYKLGFFKQYKDMMSGEGPPGABPQ 1153

## RESULT 11

US-08-476-062A-43  
; Sequence 43, Application US/08476062A  
; Patent No. 5877275

; GENERAL INFORMATION:  
; APPLICANT: Amnaut, M. Amin

; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY

; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/476,062A

; FILING DATE: 07-JUN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/216,081

; FILING DATE: 21-MAR-1994

; APPLICATION NUMBER: 07/637,830

; FILING DATE: 04-JAN-1991

; APPLICATION NUMBER: 07/539,842

; FILING DATE: 18-JUN-1990

; APPLICATION NUMBER: 07/212,573

; FILING DATE: 28-JUN-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Freeman, John W.

; REGISTRATION NUMBER: 29,066

; REFERENCE/DOCKET NUMBER: 00786/068003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1152 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

US-08-476-062A-43

Query Match 98.8%; Score 5808.5; DB 2; Length 1152;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 PNLDTENAMTFOENARGFGQSVWLOQSRVVVGAPOEIVAAORGLSYQCDYSTGSCBPI 60  
Db 17 PNLDTENAMTFOENARGFGQSVWLOQSRVVVGAPOEIVAAORGLSYQCDYSTGSCBPI 76  
Qy 61 RLQVPEAVNLSGLSLAATTPPOLLAGCPVHTCSENTYVKGCLCFGLFGLNLRQQPQK 120  
Db 77 RLQVPEAVNLSGLSLAATTPPOLLAGCPVHTCSENTYVKGCLCFGLFGLNLRQQPQK 136  
Qy 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKGELNSTIMEQLKSKTLPFLMAYSEEP 180  
Db 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTLPFLMAYSEEP 196  
Qy 181 RHFTFKFQNNPNSRLIKPTQLGLGTHATGLRKVVRELFNITNGARKNAKILFL 240  
Db 197 RHFTFKFQNNPNSRLIKPTQLGLGTHATGLRKVVRELFNITNGARKNAKILFL 256  
Qy 241 TDGKFGPGLGYEDYIPELDREGVIRYVLGFGDAFRSEKSGROELMTVASKPRPDHVFQAN 300  
Db 257 TDGKFGPGLGYEDYIPEADREGVIRYVIGVDAFRSEKSGROELMTIASKPRPDHVFQAN 316  
Qy 301 NFEALKTVQNLREKI PAIEGTCTGSSSFEHMSOEGFSAAITNSGPIILSTVGSYDWAG 360  
Db 317 NFEALKTVQNLREKI PAIEGTCTGSSSFEHMSOEGFSAAITNSGPIILSTVGSYDWAG 376  
Qy 361 GVFYLTYSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVOSLVLGAPRYQHIGLVAMPR 420  
Db 377 GVFYLTYSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVOSLVLGAPRYQHIGLVAMPR 436  
Qy 421 QNTGMWESNANVKGTQIGAYFCASICSDVDVNSGSDTLVLICAPHYETRGQVSVCP 480  
Db 437 QNTGMWESNANVKGTQIGAYFCASICSDVDVNSGSDTLVLICAPHYETRGQVSVCP 496  
Qy 481 PRGQBARWOCDAVLYGEOQPMGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRG-QARWOCDAVLYGEOQPMGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 555  
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGDLTMDGLVLTWGAQGHVLLRSQ 600  
Db 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGDLTMDGLVLTWGAQGHVLLRSQ 615  
Qy 601 PVLRVKAIKMFNPREVARNVFECNQVVKGEAGVVRVCLRVQKSTRDLREGQIQSVVT 660  
Db 616 PVLRVKAIKMFNPREVARNVFECNQVVKGEAGVVRVCLRVQKSTRDLREGQIQSVVT 675  
Qy 661 YDLALDSGRPHSRAVFNETKSTRQTVLGLTQTCETLKLQIPNCIEDPVPVLRNLP 720  
Db 676 YDLALDSGRPHSRAVFNETKSTRQTVLGLTQTCETLKLQIPNCIEDPVPVLRNLP 735  
Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALPPFKKNCNDNICODDLSITPFSMSLDCLVVG 780  
Db 736 SLVGTPLSAFGLNLRPVLAEADQRLFTALPPFKKNCNDNICODDLSITPFSMSLDCLVVG 795  
Qy 781 GPREFNVTVVRNDCGDSYRTQVTFPFLDLVSRKVSTLQNRQSRQSWRLACESASSTEV 840  
Db 796 GPREFNVTVVRNDCGDSYRTQVTFPFLDLVSRKVSTLQNRQSRQSWRLACESASSTEV 855  
Qy 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEP 900  
Db 856 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEP 915  
Qy 901 QLELPVKYAVYVWVTSHGVSCTKYLNTASNTSRVMQHOYOVNSLQORSIPISLVELVPV 960  
Db 916 QLELPVKYAVYVWVTSHGVSCTKYLNTASNTSRVMQHOYOVNSLQORSIPISLVELVPV 975  
Qy 961 RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQCDIP 1020  
Db 976 RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCORIQCDIP 1035  
Qy 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEP 1080  
Db 1036 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEP 1095  
Qy 1081 KVEPFEVNPFLPLIVGSSVGLLALLITAALYKLGFFKQYKDMMSGEGPPGABPQ 1137

Db 1096 KVEPFEVNPPLPIVGVSSVGLLLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1152  
RESULT 12  
PCT-US96-01314-43  
; Sequence 43, Application PC/TUS9601314  
; GENERAL INFORMATION:  
; APPLICANT: M. Amin Arnaout  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN  
; TITLE OF INVENTION: ANTAGONISTS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: Wordperfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/01314  
; FILING DATE: 30-JAN-96  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/380,167  
; FILING DATE: 30-JAN-95  
; ATTORNEY/AGENT INFORMATION:  
; NAME: John W. Freeman  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00786/267001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1152  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
PCT-US96-01314-43  
Query Match 98.8%; Score 5808.5; DB 5; Length 1152;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;  
Qy 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 76  
Qy 61 RLOQVPEAVNMVSLGLSLAATTSPPQLLACGPTVHQCSENTYVYKGLCFGLFGSNLRQPOK 120  
Db 77 RLOQVPEAVNMVSLGLSLAATTSPPQLLACGPTVHQCSENTYVYKGLCFGLFGSNLRQPOK 136  
Qy 121 FPALRCPCDESDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLPFLSMQYSEEF 180  
Db 137 FPALRCPCDESDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLPFLSMQYSEEF 196  
Qy 181 RIHPTFKFQNNPNRSLIKPIITQLGRHTATGLRKRVRELNIINGARKNAKFLIL 240  
Db 197 RIHPTFKFQNNPNRSLIKPIITQLGRHTATGLRKRVRELNIINGARKNAKFLIL 256  
Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVGLFGDAFSEKSRQELNTVASKPRDHDVFOAN 300  
Db 257 TDGEKFGDPLGYEDVPELDREGVIRYVGLFGDAFSEKSRQELNTVASKPRDHDVFOAN 316  
Qy 301 NFEALTKVQNLREKIPALIEGTOTGSSSPHEMSQEGFSAALITNSGILLSTVGSYDWAG 360  
Db 317 NFEALTKVQNLREKIPALIEGTOTGSSSPHEMSQEGFSAALITNSGILLSTVGSYDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436  
Qy 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYBQTRGQSVVCP 480  
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYBQTRGQSVVCP 496  
Qy 481 PRGQARWQCDVAVLYGEOQPMGRFGAALTVDLVGVDNGDKLTDVAICAGEEDNRGAVILF 540  
Db 497 PRG-PARWQCDVAVLYGEOQPMGRFGAALTVDLVGVDNGDKLTDVAICAGEEDNRGAVILF 555  
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVLTGCAQGHVLLRSQ 600  
Db 556 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLGSGQDLTMDGLVLTGCAQGHVLLRSQ 615  
Qy 601 FVLRVKAIWEPNPREVARNVFECDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 616 FVLRVKAIWEPNPREVARNVFECDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 675  
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPFVLRNF 720  
Db 676 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPFVLRNF 735  
Qy 721 SLVGTPLSAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
Db 736 SLVGTPLSAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 795  
Qy 781 GPREENVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLONQSRQSRWLACESASSTEV 840  
Db 796 GPREENVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLONQSRQSRWLACESASSTEV 855  
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNNPRTNKTEF 900  
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNNPRTNKTEF 915  
Qy 901 QLELPVKYAVVAVVTSHGVSITKYLNTASENTSRVMQHQYQVSNLQORSPLTSLVFLVEV 960  
Db 916 QLELPVKYAVVAVVTSHGVSITKYLNTASENTSRVMQHQYQVSNLQORSPLTSLVFLVEV 975  
Qy 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDPLAELKAPVNCSTAVCORIQCDIP 1020  
Db 976 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDPLAELKAPVNCSTAVCORIQCDIP 1035  
Qy 1021 PFGIOEENATLKGMLSPDWYIKTSHNHLIYVSTABILFNDSVFTLLPQGAFAVRSQET 1080  
Db 1036 PFGIOEENATLKGMLSPDWYIKTSHNHLIYVSTABILFNDSVFTLLPQGAFAVRSQET 1095  
Qy 1081 KYEPFEVNPPLPIVGVSSVGLLLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1137  
Db 1096 KYEPFEVNPPLPIVGVSSVGLLLALITAAALYKLGFFKQYKDMWSEGGPPGABPQ 1152  
RESULT 13  
5424399-2  
; Patent No. 5424399  
; APPLICANT: ARNAOUT, M. AMIN  
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/78,871  
; FILING DATE: 16-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 539,842  
; FILING DATE: 18-JUN-1990  
; APPLICATION NUMBER: 212,573  
; FILING DATE: 28-JUN-1988  
; SEQ ID NO: 2;  
; LENGTH: 1152  
5424399-2  
Query Match 98.8%; Score 5808.5; DB 6; Length 1152;

Best Local Similarity 98.6%; Pred. No. 0;									
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;									
Qy	1	FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI	76						
Qy	61	RLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRQOPQK	120						
Db	77	RLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRQOPQK	136						
Qy	121	FPALRGCPQSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSEEF	180						
Db	137	FPALRGCPQSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSEEF	196						
Qy	181	RIHFTFKFQNNPRSLIKETITOLLGRTHTATGLRKVVRELFTNGARKNAFKILFL	240						
Db	197	RIHFTFKFQNNPRSLIKETITOLLGRTHTATGLRKVVRELFTNGARKNAFKILFL	256						
Qy	241	TGGEKFGDPLGYEDVPELDREGVIRVYVVGDAFRSEKSKQELNTVASKPPRDHVFQAN	300						
Db	257	TGGEKFGDPLGYEDVPELDREGVIRVYVVGDAFRSEKSKQELNTVASKPPRDHVFQAN	316						
Qy	301	NFEALKTQNLREKIFAIEGTQSGSSSFHEMSQSGFSAITNSGPELLSTVGSYDWAG	360						
Db	317	NFEALKTQNLREKIFAIEGTQSGSSSFHEMSQSGFSAITNSGPELLSTVGSYDWAG	376						
Qy	361	GVFLYTSKSKSTFNMTRVDSMDMDAYLGVAALILNRVOSLVLGAPRYOHIGLVAMFR	420						
Db	377	GVFLYTSKSKSTFNMTRVDSMDMDAYLGVAALILNRVOSLVLGAPRYOHIGLVAMFR	436						
Qy	421	QNTGWESNANVKTQIGAYGASLCSVDVDSNGSTDLVLIGAPHYETOTRGQSVSCPL	480						
Db	437	QNTGWESNANVKTQIGAYGASLCSVDVDSNGSTDLVLIGAPHYETOTRGQSVSCPL	496						
Qy	481	PRGORARQCDAYLVGEGQDPWGRFAGALTVLGVNWDGDKLTDVAIGAPGEDNKGAYL	540						
Db	497	PRG-RARQCDAYLVGEGQDPWGRFAGALTVLGVNWDGDKLTDVAIGAPGEDNKGAYL	555						
Qy	541	HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDQDLMGDLVLTVGAGHVLILRSQ	600						
Db	556	HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDQDLMGDLVLTVGAGHVLILRSQ	615						
Qy	601	PVLRYKATMEPNPREVARNVECDQVVKGEAGEVRVCLVQKSTDRLESGOIQSVVT	660						
Db	616	PVLRYKATMEPNPREVARNVECDQVVKGEAGEVRVCLVQKSTDRLESGOIQSVVT	675						
Qy	661	YDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQPNCTEDPVPVILRLNF	720						
Db	676	YDLALDSGRPHSRVAFNETKNSRRQTVGLTQTCETLKLQPNCTEDPVPVILRLNF	735						
Qy	721	SLVGTPLSARFNLRPVLAEDAQRLEFALPPEKNGNDNICODDLSTFSPMSLDCLVVG	780						
Db	736	SLVGTPLSARFNLRPVLAEDAQRLEFALPPEKNGNDNICODDLSTFSPMSLDCLVVG	795						
Qy	781	GPRESNVTVVRNDCGDSYRTQVTFPPFDLSYRKVSTLQNRQSRWRLACESASSTEV	840						
Db	796	GPRESNVTVVRNDCGDSYRTQVTFPPFDLSYRKVSTLQNRQSRWRLACESASSTEV	855						
Qy	841	SCALKSTCSINHPIPPENSEVTENITPDVDSKASLGKLLKANKVTSNNMPNTKTEP	900						
Db	856	SCALKSTCSINHPIPPENSEVTENITPDVDSKASLGKLLKANKVTSNNMPNTKTEP	915						
Qy	901	QLELPVKAVVYVTVSHGVSTKYLNFNTASENTSRVQHOYQVSNLQORSILPISLVFLVPV	960						
Db	916	QLELPVKAVVYVTVSHGVSTKYLNFNTASENTSRVQHOYQVSNLQORSILPISLVFLVPV	975						
Qy	961	RINQTVIWRDQVTFSEMLSTCTTKRELPSHSDPLAELRKAPVNGSIAVCQRIQCDIP	1020						
Db	976	RINQTVIWRDQVTFSEMLSTCTTKRELPSHSDPLAELRKAPVNGSIAVCQRIQCDIP	1035						
Qy	1021	FFGIQEEFNATLKGSLSPDWYIKTSHNLLIVSTABILFNDSVFTLLPGQAFVRSQTET	1080						

## RESULT 14

US-08-476-062A-44  
; Sequence 44, Application US/08476062A  
; Patent No. 5877275  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY  
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,062A  
; FILING DATE: 07-JUN-1995  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 08/216,081  
; FILING DATE: 21-MAR-1994  
; APPLICATION NUMBER: 07/637,830  
; FILING DATE: 04-JAN-1991  
; APPLICATION NUMBER: 07/539,842  
; FILING DATE: 18-JUN-1990  
; APPLICATION NUMBER: 07/212,573  
; FILING DATE: 28-JUN-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00786/068003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1163 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-476-062A-44

Query Match 58.6%; Score 3446; DB 2; Length 1163;  
Best Local Similarity 60.5%; Pred. No. 4.4e-278;  
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

Qy	1	FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI	60
Db	20	FNLDTENAMTFOENARGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI	79
Qy	61	RLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRQOPQK	120
Db	80	GLQVPVEAVNMISGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPLGSLNRQOPQK	137
Qy	121	FPALRGCPQSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLFSLMQYSEEF	180
Db	136	LPVSRQECPEQEDIVFLIDGSGSIISRNFRATNMFNFRVVISQFQPSQFSLMQFSNKF	197
Qy	181	RIHFTFKFQNNPRSLIKETITOLLGRTHTATGLRKVVRELFTNGARKNAFKILFL	240

```
198 QTHFTFEERRTSNPLSLLASVHQLQCGTYTATAIQNVVHRLPHASYGARRDATKILIVI 257
241 TDGKFGDPLGYEDVTPELDREGVIRYVLGFGDAFSEKSRQELNLTVAASKPRDRHVFQAN 300
258 TUGKKGESLDYKDVIPMDAAGIRIYALGVGLAFQNRNSWKLNDIASQSEHIFKVE 317
301 NFEALKTQVQNLREKIFAIEGTQTGSSSFHEMSQEGFSAATISNGPILSTVGSYDWAQ 360
318 DFDALXDIQNLKEKIFAIEGTQTSSTSSFELEMAQEGFSAVFTPDGPVLGAVGFTWSG 377
361 GVFLYTSKEKSFIFNNRVDSWMDAYLGYAAAILRNVRVQSLVLAGPYOHI GLVAMFR 420
378 GAFLYPPNKSPTFNNQENVDMDSYLGYSTELALWKGVSQSLVLAGPYOHTGKAVIFT 437
421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEITRGQGVSCPL 480
438 QVSRQWRKAEVGTGTIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEITRGQGVSCPL 497
481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDNVNGDKLTDVWIGAPGEENRGAVYLF 540
498 PRGWR-RWMCDAVLYGEGQHPWGRFGAALTVLGDNVNGDKLTDVWIGAPGEENRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSGGGQDLMGDLVLTGGAQGHVLLRSQ 600
557 HGVLPFSISPSHSQRIAGSKLSPRLQYFGQSGGGQDLMGDLVLTGGAQGHVLLRSQ 616
601 PVLRVKAIEMFPRVARNVFECDQVVKGKAGEVTVCLHVOKSTRDLRGQTSVVT 660
617 PVLWGVSNQFTPAEIPRFAFECREQVSEQTLVQSNICLYIDKSKNLLGSRDLQSSVT 676
661 YDLALDSGSPHRAVNETKSTRRTQVLTGTCTETLKLQLPNCIEDPVPPIVLRNF 720
677 LDALDPGLSPRAFTQETKNSLSRVRLGKHCENENLLPSCVEDSVTPIRLNF 736
721 SLVGTPLSAFKNLRLVLAEDQELTALPFPKNCNDNICODDLSITFSEMSLDCLVVG 780
737 TLVGKPLAFRLRPLMAALAQRYFTASLPFEKNCQADHICQDMLG1SFSPGLKSLVG 796
781 GRPEFNVTVVRNDEGDSYRTQVTPPPLDLVSRKUSTLQNSQESQSWELCASAESTE 840
797 SNLELNAEVMWVNDGDSYRTQVTPPPLDLVSRKUSTLQNSQESQSWELCASAESTE 854
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASIGNKLKANKVTSNNMPTNKTEF 900
855 SQGTWSTSCRIHNLIFGGAQITFLATPDVSPKAVLGDRLLTANVSSNNTPRTSKTF 914
901 QLELPVKIYAVVTVSHGVSTKYLNETAS-ENTSRVMOHQYOVSNLQBSLPI SILVPLVP 959
915 QLELPVKIYAVTVSSHEQFTKYLNFSESEKSHVAMRYQVNNLQCRDLFVSI NFWVP 974
960 VRLNOTVINDRPQVTFSENLSCTCHKERLPSSHDFLAE LRKAPVNVCSIAVCORIQCDI 1019
975 VELNQEAVMDVESHVPHQNPRLCSSEKIAPASDFLAHIQKNPVLDCSIAGCLAFRCOV 1034
1020 PFFGIQEFNATLKNLSPDWIKYTSNHLITVSTAEILFNDSVFTLPGQAFVRSQTE 1079
1035 PSFSVQEEELFTLKNLSFGWRQILOKKVS VVVAEITFDTSVYSQLPQGEAFWEAQT 1094
1080 TKVBEPEVNPPLVINGSSVGLLALLALITAYLKGPFKRYKQVDMWSE 1128
1095 TVLEKTKVHNPLPLVINGSSIGGLLALLITAVLYKVGFFKRYKQVDMWSE 1143
```

## RESULT 15

PCT-US96-01314-44

; Sequence 44, Application PC/TUS9601314

; GENERAL INFORMATION:

; APPLICANT: M. Amin Attacut

; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN

; TITLE OF INVENTION: ANTAGONISTS

; NUMBER OF SEQUENCES: 78

; CORRESPONDENCE ADDRESS:

```
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US96-01314-44
```

Query Match 58.6%; Score 3446; DB 5; Length 1163;

Best Local Similarity 60.5%; Pred. No. 4.4e-278;

Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

```
QY 1 FNLDTENAMTQENARGFGQSVWOLQGSRVVVGAPQEI VAAQSGSLYQCDYSTGSCPEI 60
DB 20 FNLDTTELTA RVDSAGFGDSVQVYANSVWVGAPQKITAANQTGGLYQCGYSTGACEPI 79
QY 61 RLQVPEAVNMSLGLSLAATSPPELLACGTVHQTSSENTYVKGLCPLFGSNLRQOQOK 120
DB 80 GLQVPPPEAVNMSLGLSLASTTSPQLLACGPTVHHCORNNYLTGLCFLAFT -QLTOR 137
QY 121 FPEALRGCPQSDSIAFLIDGSGSII PHDFRMKELVSTIMEQLKCKSTLFLSLMOYSEEF 180
DB 138 LPVSRQECPRQEDIVELIDGSGSISRNFAFMNFRAVISQRPSTQSLMQFSNKP 197
QY 181 RIHFTPEFQNNPNRSLIKFITQLLGRHTHTATGLRKVYRELFTNGARKNAFKILFL 240
DB 198 QTHFTFEFRRTSNPLSLLASVHQLQCGTYTATAIQNVVHRLPHASYGARRDATKILIVI 257
QY 241 TDGKFGDPLGYEDVTPELDREGVIRYVLGFGDAFSEKSRQELNLTVAASKPRDRHVFQAN 300
DB 258 TDGKKGESLDYKDVIPMDAAGIRIYALGVGLAFQNRNSWKLNDIASQSEHIFKVE 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTGSSSFHEMSQEGFSAATISNGPILSTVGSYDWAQ 360
DB 318 DFDALXDIQNLKEKIFAIEGTQTSSTSSFELEMAQEGFSAVFTPDGPVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSFIFNNRVDSWMDAYLGYAAAILRNVRVQSLVLAGPYOHI GLVAMFR 420
DB 378 GAFLYPPNKSPTFNNQENVDMDSYLGYSTELALWKGVSQSLVLAGPYOHTGKAVIFT 437
QY 421 QNTGHWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEITRGQGVSCPL 480
DB 438 QVSRQWRKAEVGTGTIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEITRGQGVSCPL 497
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDNVNGDKLTDVWIGAPGEENRGAVYLF 540
DB 498 PRGWR-RWMCDAVLYGEGQHPWGRFGAALTVLGDNVNGDKLTDVWIGAPGEENRGAVYLF 556
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 17:05:16 ; Search time 14.559 Seconds  
(without alignment)  
7512.163 Million cell updates/sec

Title: US-09-902-481B-4  
Perfect score: 5884  
Sequence: 1 FNLDTENAMTQENARGFGQ.....FRQYKDMWSEGGPRGAPQ 1137

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1 RWHU1B	Cell surface glyco
2	4460	75.8	1153	2 S00551	leukocyte surface
3	3469	59.0	1163	1 RWHU1C	cell surface glyco
4	1543.5	26.2	1170	2 S03308	cell surface glyco
5	1527.5	26.0	1163	2 I56126	lymphocyte fuction
6	1140	19.4	1179	2 A53213	integrin alpha-E c
7	1093.5	18.6	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45914	integrin alpha 2 s
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1055	17.9	1180	2 A35854	integrin alpha-1 c
11	1054	17.9	1181	2 A33998	integrin alpha-2 c
12	663	11.3	1039	2 R41131	lymphocyte-Peyer's
13	642	10.9	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.4	1041	2 T31437	integrin alpha cha
16	579.5	9.8	1054	2 JC7294	alphanp integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.6	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.4	1053	2 S44350	integrin alpha-5 c
20	546.5	9.3	1034	2 A36108	integrin alpha-V c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	534.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	fibronectin recept
24	532	9.0	1073	2 B36429	integrin alpha-6 c
25	530.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	8.9	1048	2 A27431	integrin alpha-5 c
27	525.5	8.9	1091	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.4	1146	2 S40311	integrin - fruit f

## ALIGNMENTS

### RESULT 1

#### RWHU1B

Cell surface glycoprotein CD11b precursor [validated] - human  
N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein M  
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000  
C:Accession: A31108; A28915; A30892; A32218; A46526; A26091; I52567  
R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.  
J. Biol. Chem. 263, 12403-12411, 1988  
A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C  
B.  
A:Reference number: A31108; MUID:88315033; PMID:2457584  
A:Accession: A31108  
A:Molecule type: mRNA  
A:Residues: 1-1153 <COR>  
A:Cross-references: GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148  
A:Note: part of this sequence was confirmed by protein sequencing  
R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.  
J. Cell Biol. 106, 2153-2158, 1988  
A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor  
A:Reference number: A28915; MUID:88257215; PMID:2454931  
A:Accession: A28915  
A:Molecule type: mRNA  
A:Residues: 1-499,501-965, 'P', 967-1153 <ARN>  
A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA59  
A:Note: The authors translated the codon TAC for residue 1129 as Thr  
A:Note: part of this sequence, including the amino end of the mature protein, was confi  
R:Shelley, C.S.; Arnaout, M.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991  
A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally re  
A:Reference number: A41600; MUID:92073318; PMID:1683702  
A:Accession: A41600  
A:Molecule type: DNA  
A:Residues: 1-9 <SHE>  
A:Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215  
R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.  
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988  
A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhes  
A:Reference number: A94193; MUID:88190151; PMID:2833753  
A:Accession: A30892  
A:Molecule type: mRNA  
A:Residues: 917-1042 <AR2>  
A:Cross-references: GB:M18044  
R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.  
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989  
A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence recepto  
A:Reference number: A32218; MUID:89098893; PMID:2563162  
A:Accession: A32218  
A:Molecule type: mRNA  
A:Residues: 9-1153 <HIC>  
A:Cross-references: GB:J04145; NID:g189068; PIDN:AAA59903.1; PID:g386975

30 495.5 8.4 1039 2 A34269 integrin alpha-2b  
31 494 8.4 1394 2 A29637 position-specific  
32 489 8.3 1137 2 JC5950 integrin alpha-7 c  
33 488 8.3 1135 2 I61186 alpha-7 integrin -  
34 486.5 8.3 1037 2 A60163 glycoprotein IIb -  
35 486 8.3 126 2 B30892 leukocyte adhesion  
36 470 8.0 1106 2 S38783 integrin alpha cha  
37 462 7.9 1226 2 S44824 F54P2.1 protein -  
38 453 7.7 1045 2 S60571 integrin alpha v c  
39 444.5 7.6 1139 2 S28277 hypothetical prote  
40 426 7.2 1115 2 T09403 integrin alpha cha  
41 421.5 7.2 1115 2 T09433 glycoprotein IIb -  
42 391 6.6 764 2 I36916 integrin alpha cha  
43 309.5 5.3 1086 2 I18523 glycoprotein IIb -  
44 299 5.1 604 2 I36917 integrin alpha-1 -  
45 286.5 4.9 272 2 A55348 integrin alpha-1 -

A;Note: part of this sequence was confirmed by protein sequencing  
 R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.  
 J. Immunol. 150, 480-490, 1993  
 A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in  
 during evolution.  
 A;Reference number: A46526; MUID:93123748; PMID:8419480  
 A;Accession: A46526  
 A;Status: not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 1-499, 501-1153 <FL>  
 A;Cross-references: GB:552227; NID:9263047; PIDN:AA24821.1; PID:G263049  
 A;Note: the last three bases of intron 13, CAG, are included in some but not all mature  
 A;Note: sequence extracted from NCB1 backbone (NCB1P:121963)  
 R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaut, M.A.  
 Biochim. Biophys. Acta 874, 368-371, 1986  
 A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp  
 A;Reference number: A90664; MUID:87076671; PMID:3539202  
 A;Accession: A26091  
 A;Molecule type: protein  
 A;Residues: 17-31 <PI>  
 A;Experimental source: granulocytes  
 R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.  
 Blood 79, 865-870, 1992  
 A;Title: Characterization of the myeloid-specific CD11b promoter.  
 A;Reference number: 152567; MUID:92144986; PMID:1346576  
 A;Accession: 152567  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-9 <RES>  
 A;Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:G553219  
 C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
 C;Genetics:  
 A;Gene: GDB:ITGAM; CR3A  
 A;Cross-references: GDB:120599; OMIM:120980  
 A;Map position: 16p11.2-16p11.2  
 C;Superfamily: cell surface glycoprotein Cblb; von Willebrand factor type A repeat hom  
 C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag  
 P:1-16/Domain: signal sequence #status predicted <SIG>  
 P:17-1153/Product: cell surface glycoprotein CD11b #status experimental <NAT>  
 P:17-1108/Domain: extracellular #status predicted <EXT>  
 P:148-318/Domain: von Willebrand factor type A repeat homology <VWR2>  
 P:465-473/Region: calcium/magnesium binding #status predicted  
 P:530-538/Region: calcium/magnesium binding #status predicted  
 P:593-601/Region: calcium/magnesium binding #status predicted  
 P:1109-1134/Domain: transmembrane #status predicted <TM>  
 P:1135-1153/Domain: intracellular #status predicted <INT>  
 P:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.5%; Score 5852; DB 1; Length 1153;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTEAMTFOENARGFGQSVVQLQGSRRVVGAPQETVAANQSGSLYQCDYSTGSCPEI 60  
 DB 17 FNLDTERAMTFOENARGFGQSVVQLQGSRRVVGAPQETVAANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVKGCLFGLGSMRLRQOPQK 120  
 DB 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVKGCLFGLGSMRLRQOPQK 136

QY 121 PPEALRGCPQEDSDIAPLIDGSGSLIIPHDPRMKEWSTVMEQLKSKTLFSLMQYSEEF 180  
 DB 137 PPEALRGCPQEDSDIAPLIDGSGSLIIPHDPRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196

QY 181 RIHFTFEFQNNPNSRLIKEITQLLGRTHATGLRKVYRELFINITGARKNAFKILFL 240  
 DB 197 RIHFTFEFQNNPNSRLIAKVEITQLLGRTHATGLRKVYRELFINITGARKNAFKILV 256

QY 241 TDGKFKGDPGLGYEDVPELDEGVIRVYVGVDAFRKSKHQLNTVASKPPRDHVFQIN 300  
 DB 257 TDGKFKGDPGLGYEDVPEADREGVIRVYVGVDAFRSEKSKHQLNTASKPPRDHVFQVN 316

QY 301 NFEALKTIONQUREKIPFAIEGTQTGSSSFEHMQEGRSAATNSNGPILLSVTGSDWAG 360  
 DB 317 NFEALKTIONQUREKIPFAIEGTQTGSSSFEHMQEGRSAATNSNGPILLSVTGSDWAG 376

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNVRQSLVLCAPYQHIGLVAMFR 420  
 DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNVRQSLVLCAPYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDNGSGTDLVIGAPHYYEOTRGQNSVCPL 480  
 DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDNGSGTDLVIGAPHYYEOTRGQNSVCPL 496

QY 481 PRGQARAWQCDVAVLGEQOQFGRFGAALTVDGVNGDKLTDVAIGAPCEEDNRGAVYLP 540  
 DB 497 PRGQARAWQCDVAVLGEQOQFGRFGAALTVDGVNGDKLTDVAIGAPCEEDNRGAVYLP 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTVGAQGHVLLRSQ 600  
 DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTVGAQGHVLLRSQ 616

QY 601 PVLRYKAIWEPNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
 DB 617 PVLRYKAIWEPNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

QY 661 YDLALDSGRPHSRAVENETKSTRQTQVGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720  
 DB 677 YDLALDSGRPHSRAVENETKSTRQTQVGLTQTCETLKLQLPNCIEDPVSPVILRLNF 736

QY 721 SLVGTPLSAFQNLRLPVLAEDQRLFTALPPEKNGNDNICODDLSITSEKSLDCLVWG 780  
 DB 737 SLVGTPLSAFQNLRLPVLAEDQRLFTALPPEKNGNDNICODDLSITSEKSLDCLVWG 796

QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVYSTLQNRQSRWSLACESASSTEV 840  
 DB 797 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVYSTLQNRQSRWSLACESASSTEV 856

QY 841 SGALKSTCSINHPIPPENSEVTNFTTFVDSKASLGNKLLKANTSENMPRTNKTSE 900  
 DB 857 SGALKSTCSINHPIPPENSEVTNFTTFVDSKASLGNKLLKANTSENMPRTNKTSE 916

QY 901 QLLELPKYAVVMVVTSHGVTYKLNFTASENTRVNHQYQVSNLQSRSLPISLVLPV 960  
 DB 917 QLLELPKYAVVMVVTSHGVTYKLNFTASENTRVNHQYQVSNLQSRSLPISLVLPV 976

QY 961 RLMTQVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCSAVCORICDIP 1020  
 DB 977 RLMTQVIMDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCSAVCORICDIP 1036

QY 1021 FFGIQEEFNATLKGNLSPDWYIKTSHNHLIYSTAEILFNDSVFTLLPGQGFVRSQTET 1080  
 DB 1037 FFGIQEEFNATLKGNLSPDWYIKTSHNHLIYSTAEILFNDSVFTLLPGQGFVRSQTET 1096

QY 1081 KVEPPEVPNPPLPLTVGSSVGGILLALITAAALYKLGFFKRYQKDMMEGPPGABPQ 1137  
 DB 1097 KVEPPEVPNPPLPLTVGSSVGGILLALITAAALYKLGFFKRYQKDMMEGPPGABPQ 1153

RESULT 2  
 S00551  
 leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse  
 N;Alternate names: complement-3 receptor alpha chain  
 C;Species: Mus musculus (house mouse)  
 C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 22-Oct-1999  
 C;Accession: S00551; 159078  
 R;Pytel, K.  
 EMBO J. 7, 1371-1378, 1988  
 A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the  
 A;Reference number: S00551; MUID:88312584; PMID:3044779  
 A;Accession: S00551  
 A;Molecule type: DNA  
 A;Residues: 1-1153 <P>  
 A;Cross-references: EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G52983  
 A;Note: the authors translated the codon CAC for residue 569 as Gin

R:Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986  
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor  
A:Reference number: 159078; MUID:86287312; PMID:2942940  
A:Accession: 159078  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 11-44 <RES>  
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193  
C:Genetics:  
C:Gene: Mac-1  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology  
C:Keywords: cell adhesion; glycoprotein; transmembrane protein  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental  
F:148-319/Domain: von Willebrand factor type A repeat homology <VWA2>  
F:1106-1129/Domain: transmembrane #status predicted <TM>  
  
Query Match 75.8%; Score 4460; DB 2; Length 1153;  
Best Local Similarity 73.8%; Pred. No. 66-304;  
Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;  
  
QY 1 FNLDTENAMTFOENARGGOSVVOLOGSHVVVGAPQEIIVAAHQSGLYOCDYSTGSCBPI 60  
DB 17 FNLDTEHPMTFOENARGGQNNVQLGGTSVVVAARQEAQKVVMTGALYQCDYSTGRCHI 76  
  
QY 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVXGLCFGLFGLNLRQOPK 120  
DB 77 PLQVPEAVNMSLGLSLAVSTVPQQLACGPTVHQNCKENTYVNGLYGLFGLNLRPQQ 136  
  
QY 121 PPEALRGCOEDSDIAFLDGGSSIIIPDFRMKKEVSTVMQLKSKTLPFLMOYSREB 180  
DB 137 PPEALRECPQESDIIVFLDGGSSINIDFQKMKFPVSTMQFKSKTLPFLMOYSDEP 196  
  
QY 181 RIHFTKPEFONNPRLSKPIPTQLGRTHATGLKVVRELPENITNGARKNAKIFLIL 240  
DB 197 RIHFTNDKRNPSRSHVSPKQLNGRIKTASGIRKVVRELPFKTNGARENAKILVVI 256  
  
QY 241 TDGEKFGDPLGYEDVITPELDREGVIRYVIGVGDAFRSEKSRRELNTVASKPRDHWQIN 300  
DB 257 TDGEKFGDPLGYEDVITPELDREGVIRYVIGVGDAFRSEKSRRELNTVASKPRDHWQIN 316  
  
QY 301 NPEALATIQNLQEKIPIAEGTGTGSSSFEHMSQEGFSAAITSGPLISTVGSVDWAG 360  
DB 317 NPEALATIQNLQEKIPIAEGTGTGSSSFEHMSQEGFSAAITSGPLISTVGSVDWAG 376  
  
QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAAIIILNRNVQSLVVLGAPRYOHIGLVMPR 420  
DB 377 GAFLYTSKSKSTFINMTRVDSMDNDAYLGVAAIIILNRNVQSLVVLGAPRYOHIGLVMPR 436  
  
QY 421 QNTGMESNANVKTGTGAYFGASLCSVDVDSNGSTDVLILGAPHYVEOTRGQSVSCPL 480  
DB 437 ENFGTWPHTSIRKSGISGTSYFGASLCSVDVDSNGSTDVLILGAPHYVEOTRGQSVSCPL 496  
  
QY 481 PRGORARWQCDVILYEGQGFWRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYILF 540  
DB 497 PRG-RARWQCDVILYEGQGFWRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAYILF 555  
  
QY 541 HTSGSGISPSHQRINGSKLSPRLQYFGOSLGGQDLTMDGLVDITVGAQCHVALLRSQ 600  
DB 556 YGASIASLSASHSHRIIAGHSPGLQYFGOSLGGQDLTMDGLVDITVGAQCHVALLRSQ 615  
  
QY 601 PVLVRKAIEMFNPREVARNVFCNDQVVKGEVRVCLVHVKSTRDLRREGQIOSVVT 660  
DB 616 PVLRLKAIEMFNPREVARNVFCNDQVVKGEVRVCLVHVKSTRDLRREGQIOSVVT 675  
  
QY 661 YDLALDGRPHSRVAVNETKNTSTRQTVGLQTCETLKLQPLNCIEDVPSIVLRINP 720  
DB 676 YDLALDGRPHSRVAVNETKNTSTRQTVGLQTCETLKLQPLNCIEDVPSIVLRINP 735  
  
QY 721 SLVCTPLSARCNRPVLAEDAQRILFTALPPEKNCNDNICODDLSTTFSFMSLDCILVVG 780  
DB 736 TLVGEPLSRPNCNRPVLAEDAQRILFTALPPEKNCNDNICODDLSTTFSFMSLDCILVVG 795

QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQRSWRL-ACESASSTE 839  
DB 796 GPQDFNVSTLRNDGEDSYRTQVTVYPSGLSYRKDSASQNPITKPFVFKPAESSSSSE 855  
  
QY 840 VSGALKSTSCSINHPIPEPENSEVTENITPDVDSKASGKLLKANKVTSNNMPTNKTE 899  
DB 856 GHGALKSTTWNHPIPEPENSEVTENITPDVDSKASGKLLKANKVTSNNMPTNKTE 915  
  
QY 900 FOELPLVKYAVVTVTSHGVSTKYLNTFASNTSRVNMHQYQVSNLQORSLPISLVLFP 959  
DB 916 FOELPLVKYALVTVTSHGVSTKYLNTFASNTSRVNMHQYQVSNLQORSLPISLVLFP 975  
  
QY 960 VRLNQTVMWRPQVTFPBNLSSTCHTKERLPSHSDFLAELKAPVNCSTAVCORICODI 1019  
DB 976 VQINNVTVMWRPQVTFPBNLSSTCHTKERLPSHSDFLAELKAPVNCSTAVCORICODI 1035  
  
QY 1020 PFGIGQEENATLKGNSLSPDWYIKTSHNHLIYSTABILFNDVSFTILLPGQGAFFRSQTE 1079  
DB 1036 PFWTQIRLNFVTLKGNLSPDWYIKTSHNHLIYSTABILFNDVSFTILLPGQGAFFRSQTE 1095  
  
QY 1080 TKVPPFVFNPLPLIVGSSVGGILLALLITANLYKLGFFKQYKDMSEGGPGGAPQ 1137  
DB 1096 TKVPPFVFNPLPLIVGSSVGGILLALLITANLYKLGFFKQYKDMSEGGPGGAPQ 1153  
  
RESULT 3  
RHEUIC  
cell surface glycoprotein CD11c precursor - human  
C:Alternate names: leukocyte adhesion receptor p150,95 alpha chain  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text\_change 22-Jun-1999  
C:Accession: A36584; A35543; S00864  
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
J. Biol. Chem. 265, 12750-12751, 1990  
A:Reference number: A36584  
A:Contents: erratum  
A:Accession: A36584  
A:Molecule type: DNA  
A:Residues: 1-1163 <COR>  
A:Note: this revision to the sequence from reference A35543 includes the carboxyl end  
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
J. Biol. Chem. 265, 2782-2788, 1990  
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.  
A:Reference number: A35543; MUID:90153906; PMID:2303426  
A:Accession: A35543  
A:Molecule type: DNA  
A:Residues: 1-834 <CO2>  
A:Note: this sequence has been revised in reference A36584  
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.  
EMBO J. 6, 4023-4028, 1987  
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte  
A:Reference number: S00864; MUID:88166645; PMID:3327687  
A:Accession: S00864  
A:Molecule type: mRNA  
A:Residues: 1-755, 'L', 757-1163 <CO3>  
A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830  
A:Note: part of this sequence was confirmed by protein sequencing  
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my  
C:Genetics:  
A:Gene: ITGAX; CD11C  
A:Cross-references: GDB:119758; OMIM:151510  
A:Map position: 16p11.2-16p11.2  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homol  
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat,  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>  
F:20-1107/Domain: extracellular #status predicted <EXT>  
F:149-319/Domain: von Willebrand factor type A repeat homology <VWA4>  
F:1108-1163/Domain: transmembrane #status predicted <TM>  
F:1134-1163/Domain: intracellular #status predicted <INT>  
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match	59.0%	Score 3469	DB 1	Length 1163
Best Local Similarity	60.9%	Pred. No. 1.Be-293		
Matches	688	Conservative 142	Mismatches 234	Indels 6
		Gaps		
Qy	1	FNLDTENAMTFQENARCGGVSVOLOQSGRVVVGAPQIBVAANRGSLYQCDYSTGSCPEI	60	
Db	20	FNLDTBELTAFRVDSAGFGDSVQVYANSWVVGAPQKIITAANTGGLYQGYSTGACEPI	79	
Qy	61	RLQVPEAVNMSLGLSLAAATSPPOLLACGFTVHOTCSENTYVKGLCFPLPGSNLRQOPK	120	
Db	80	GLQVPPPEAVNMSLGLSLASTTSPSQLACGFTVHBCGRNMYLTGLCFLGLPT--QUTQ	137	
Qy	121	FPBALRCPCQSDSIAFLIDGSGSIIPHDRFRKKEWSTVMEOLKKSKTLFSLMWQSEEF	180	
Db	138	LPVSRQCPCEQODIVFLIDQSGSISRNFAITMWNFPAVISOQRPSTQFSLMQSNKF	197	
Qy	181	RIHFTPEFQNNPNRSLIKPIITOLLGRTHATGLRKYVRELFNITNGARNAFKILFL	240	
Db	198	QTHFTPEEFARTSNPLSLLASVHOLQGTVTATAIQWVHRLFHASYGARDATKILIVI	257	
Qy	241	TDGEKFGDPLGYENVITPELDREGVIRVVGVDGAFSEKSEKQELHNTVAKPPDRHVPQIN	300	
Db	258	TDGKKGEGSLOYKDVIPMAADGIIRVIAVGVLAFQNRNWKSEINDASKPSQEHFKVE	317	
Qy	301	NFEALKTIQNLREKIFAIBGTOTGSSSSSEHEMSQEGFSAATNSGILLSTTVGSDWAG	360	
Db	318	DFDALKDQIQNLKEKIFAIEGTETTSSSFELEMAQEGFSAVTPDGPVLGAVGSPFMSG	377	
Qy	361	GVFLYTSKSKSTFNNTRVDSMDKDAYLGYAAAILNRVOSIVLGNAPRVOHGLVAMFR	420	
Db	378	GAFLYPPNMSPTFNNMQENVDMEDSYLGYSTELALMKGVQSVLGNAPRYOHTGKAVIT	437	
Qy	421	QNTGWESNANVKGTQIGAVFGSLCSVDVDSNGSTDLVILGAPHYYEQTGQGVSVCP	480	
Db	438	QVSRQWEMKAEVTGTQIGSVFGSLCSVDVDTGSDTLVILGAPHYYEQTGQGVSVCP	497	
Qy	481	PRGORARWQCDAVILYGGQGPWGFAGALTNLGVNNGDKLTDVAIGAPGEDNRGAYILF	540	
Db	498	PRGWR--RWNCDAVILYGGQHPWGFAGALTNLGVNNGDKLTDVVGAPGEENRGAYILF	556	
Qy	541	HGTSGGTSPSHSQRIAGSKLSRLQVFGSLSGQDLTMDGLVDLTWGAQGHVLLRSQ	600	
Db	557	HGVLPSTSPSHSQRIAGSKLSRLQVFGQALSGQDLTDGLVDLAVGARGQVLLATR	616	
Qy	601	PVLRYKAIMFNPREVARNFECNDQVVKGEAGEVAVLHVQKSTDRRLREGIQSVVT	660	
Db	617	PVLWVGVMQMFIPAEIPSAFECREQVVSEQTLVQSNICLYDKRSKNLGSRLQSSVT	676	
Qy	661	YDLALDSGRPHSAVFNETHKSTRQVOTGLTQTCETLKLQPCNIEDPVSPIVRLNPF	720	
Db	677	LDLALDPGRLSPRATFOETKRSLSRVKVLGLKXHCENFNLLPSCVEDSVPTILRLNF	736	
Qy	721	SLVETPLSAFQNLPRVLAEDAQRLFTALFPPEKNCNNDNICODDLSTTFPSMSLDCLVVG	780	
Db	737	TEVGKPLLAFLNLRPMLAADQAQRYFTASLPEKKNCGADHICQDNLGIFSFSPGGLKSLVG	796	
Qy	781	GPRFNVTVTVRNGEDSYRTQVTFPPPLDLSYKSVSTLQVRSQRORWELACESASTEV	840	
Db	797	SNLEINAEVMVWNGEDSYGTITFSPHAGLSYRYVAEGQKQGLRSILHITCDSPAGV--	854	
Qy	841	SGALKSTCSNHPFPENSEVTENITFDVDSKASLGNKILKXANVTSENMMPTNKTTF	900	
Db	855	SGQTWSTSCRINHLIFRGAQITFLATFDVSPKAVLGDRLLTANVSENNTPTSKTTP	914	
Qy	901	QLELPVKYAVTVWTVSHGVSTKYLNFITAS--ENTGRVMQHQVQSNLQORSPLPTSLVFLVP	959	
Db	915	QLELPVKYAVTVWVSSHEQFTKYLNFSESEKESHVAMHRYQVNVNLQORDLPVSNFVWP	974	
Qy	960	VLNQVTVLDRPQVTFPSNLSASTCHTKERLPSHSDPLAELFKAPVWNCSTIAVCORICDI	1019	
Db	975	VELNQEAVMVDVEVSHQNPQSLRCSSEKIAIPADFDLFIHQKPNVLOCSLACLRFRCDV	1034	
Qy	1020	PFFGIQEBFNATLKGNSLFDWIKYITSHNHLITVSTAEILFVNDSVFTLLPGQAGFVRSQTE	1079	



QY 768 TFSFMSLDCLVVGSP-----RBFNVTVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQV 821  
DB 780 SSPARS-----GFLRLMSSASLAVENTLNSGDAYVRLDLDFRGLSFRKVEMLQ- 831  
QY 822 QRSQRWRLACASASTEVSGAL-KSTSCSINHPIIPENSEVTNITFDVDSKASLGNKL 880  
DB 832 --PHSRMPVSCBEL--TEGSSLLTKLKNVSPFKAGQGVSLQVFNFTLLNSSWEDFV 887  
QY 881 LLKANTSEN-NMPTNKTFFQLELPVKYAVYVWVTSHTVSTYKLNFTASENTSRVWQH 939  
DB 888 ELNGTVHCENENSLOQDSNAAATHIPVLPVNLTKQENSTLYISFTPKGPTQVQHV 947  
QY 940 YQVSNLQORSPLISLVLFPVRLNQTVIDRPO-----VTFSENLS--TCHTKE-RLP 990  
DB 948 YQV-----RIQPSAYDENMET-LEALVGVPRPHSEDLITVTSVQTDPLVTCHEDLKRP 1001  
QY 991 SHSDFLAELKAPVNVCSIAVCORIQDIPFGIOBEFNATLKNLSFDWYIKTSHNLL 1050  
DB 1002 SSE--AEQCLPGV-----QFRCPIVP---RWELLQVGTVELSKEIKAS-STLS 1046  
QY 1051 IVSTAELFNDSTVTLPGQAFVRSQTETKVPPFVFNPLPIVIGSSVGLLLALITA 1110  
DB 1047 LCSLSVSFNSSKHFHLYGSKA-SEAQVLKVDLHKEKMLHYVVLSGIGGLVLLFLIFL 1105  
QY 1111 ALYKLGFEKQYKDM-SEGPPGAP 1136  
DB 1106 ALYKVGFFKRLKEKMEADGGVNGSP 1132

RESULT 6  
A:Species: Homo sapiens (man)  
C:Date: 19-Oct-1995 #sequence\_revision 31-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: A53213  
J:Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.  
J. Biol. Chem. 269, 6016-6025, 1994  
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (B) subunit. UN  
A:Reference number: A53213; MUID:94164962; PMID:8119947  
A:Accession: A53213  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1179 <SHA>  
A:Cross-references: GB:L25851; NID:G457244; PID:G457245  
C:Genetics:  
A:Gene: GDB:ITGAE  
A:Cross-references: GDB:330801  
A:Map position: 17p13  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
F:193-371/domain: von Willebrand factor type A repeat homology <WMA3>

Query Match 19.4%; Score 1140; DB 2; Length 1179;  
Best Local Similarity 28.9%; Pred. No. 3.5e-71;  
Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSLYQCDYSTGS--CEPT-RLQVP-----VEAVNSLGLSLAATTPOLLACGPVHQ 95  
DB 65 GPLHRCSLVQDEILCHPEHVPIPKGRHGVTVVRSHHGLICI-----QVLRRP--HS 117  
QY 96 TCSENTYVKGCLFLPGSNLRQOPQ----- 119  
DB 118 LSSELT---GTCSLGLPDLRFPQANFPDENLDPDARVDVTGDCYSNKGSGGDDVNTA 174  
QY 120 KFPALRGCPQED-----SDIAFLDGSGLIIPHPFRMKENVSTVNEQL-- 164  
DB 175 QRRALEKEEEDKEEEDDEEBAETAIILDGSGLIDPDPQAKDFISNNMRNRYE 234  
QY 165 KSKTSLRSLMYSSEFRHFFPKFONNPNRSLIKETOLLGTHATGTLGRKVVRLFN 224  
DB 235 KCFCNFAVQYGVQIFEDRLSDQVMSIARVQNTIQGSVTKKASAMQHVLDISFT 294  
QY 225 ITNGARKNAKILFLLTIGKEFGDPLGYEDVIPEDREGVIRYVIGVDAFRSEKSQL 284

## RESULT 7

A45226

integrin alpha-1 chain - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Sep-2003

C:Accession: A45226

R:Briesewitz, R.; Bostein, M.R.; Marcantonio, E.E.

J. Biol. Chem. 268, 2989-2996, 1993

A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit  
A:Reference number: A45226; MUID:93155124; PMID:8428973

DB 295 SSHGRRKASKVMVVLTDGIFEDPLNLTVTVINSPKMGQVERFAIGVGEFPKSAKTAREL 354  
QY 285 NTVASKPRDRHVQINNFEALKTIOQLREKIFAIEGTQTGSSSSSEHEMSQEGSAAIT 344  
DB 355 NLIADDPDETHAFKVTNYALDGLSKLYNIISMEGT---VGDALHYOLAQIGFSAQIL 411  
QY 345 SNGP--LLSTVGSYDNAGGVFLY-TSEKSTTINMTVRDSDMDA-----YLGAAAILLEN 398  
DB 412 DERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTFAAAAADABAAQCYSLGVAVAVLHKT 471  
QY 399 RVQSLVIGAPRYOHLGLVAMER-QNTGMWSENANY-KGTQIGAYFGASICSVDVDSNGST 456  
DB 472 CSLSVAGAPQKHEG--AVFELQKEGREASFLPVLEGEOMGSIYFSELCPOVIDMDGST 529  
QY 457 DLVLICAPHYEQTRGGQVSVCPFRQORARWQCDVILYGSQCGQWGFAGALTIVLGDVN 516  
DB 530 DFLVAAAPFHVHGEGRVYVYVRLSE-QDGSFSLARILISGHPGFTNARFGFAMAAGDUS 588  
QY 517 GPKLTDVAIGAP---GEEDNR--CAVYLFHGTSGGSPSHSQRISAGSKLSPLRQYFQ 570  
DB 589 QDKLTDVAIGAPLEGFADGASFGSVIYNG-HWDGLSASPSQRIASTVAPGLQYFGM 647  
QY 571 SLSGQDQLTMDGLVDLTVGAQGHVLLRSQPVLRVKALIMFNPRVARNVFECNDQVVG 630  
DB 648 SWAGGFDISGDLADITVGTLCQAVVFRSRPVVRLKVSMAFTPSALP-----IGF 697  
QY 631 KEAGEVYVCLHVQKSTRDLREGQ-QSVYVYDLALDSGRPHSRVAFNETKNTREQTQVL 690  
DB 698 NGVAVNRLCFEI-SSVTTASBSGLREALNPLFDVVGKQRRRLQCSVRSCLGCLREWS 756  
QY 691 GLTQTCETLKLQLPN---CIEDPVSPVILNLFSLVGTPLSAFGNLRPLVLAEDAQRLFT 746  
DB 757 SCSQLCEDL-LMPTEGELCEDCFSNASVKVSYQL-OTPEGQTOHPQILDRYTEPPAI 814  
QY 747 ALPPEKNGCNDNI CODDLISITFSFMSLDCLVVGPRBNVTVTVRNDGDSYRTQVTFP 806  
DB 815 FOLPVEKACKNLCFVAELQLA-TTVSQELVVGUTKELTMINETNSGDSYMTSMALN 873  
QY 807 FPLDLSYRKVSTLQORSORSLACESASSTEVSGALKSTSCSINHPIIPENSEVTNFI 866  
DB 874 YPRNLQ-----LKMQKPPSPNIQCDDPQPV---ASVLIIMNCRIGHPLV-KRSSAHVS 923  
QY 867 TFDVDSKASLGNKLKLLKANVTSENN-----MPTNKTBEQ---LELPKVAVYVWVTSHV 919  
DB 924 VMQLEENAFNRTADITVTVTNSNERRSLANETHLQFPHGVAVLSKPSIMYVNTGQGL 983  
QY 920 S--TKYLNFTASENTSRVWQHQVSNLQORSPLISLVLFPVRLNQTVIDRPOVTFSE 977  
DB 984 SHHKEFLFHVHGEN---LFGASYQ-----LQICVPTKLRLGLQVAAVKLRTQ 1028  
QY 978 NLSSTCHTKERLPSSHSDFLAELRKAPVNVCSIAVCORIQDIPFGIOBEFNATLKNLS 1037  
DB 1029 ASTVCTWISOERACAYSS-VQHVVEENHVSVCVIA-----SKENVTVAABIS 1073  
QY 1038 EDWYIKTSHNLLIVST-----AEILFNDSTVTLPGQAFVRSQETKVPFEPVNPPL 1091  
DB 1074 WD-----HSELLKDVLELQILGELISFNKSLYEGNAENH--RTKITVTVFLKDKYHSL 1125  
QY 1092 PLTVGSSVGGLLALLALITAAALYKLGFFKQYKDMSE 1128  
DB 1126 PIIKISGVGGLLVILVILFKCGFFKRYQQLANL 1162

A:Accession: A45226  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1151 <BRI>  
A:Experimental source: hepatoblastoma cell line HepG2  
A>Note: sequence extracted from NCBI backbone (NCBIP:124326)  
F:142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match 18.6%; Score 1093.5; DB 2; Length 1151;  
Best Local Similarity 27.8%; Pred. NO. 6.1e-68;  
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

Qy 1 FNLDTENAMTFQENARG-FQGSVVQL---QGSRVVVGAPQEIIVAAQNRGLSQCCDYSTGS 56  
Db 1 FNVVGNKSMTFSGPVEDMGYTVQYENEEGKWLIGSPVAGPKNRTGDVYKCPVGRGE 60

Qy 57 CSP-ILQVQVRA-----VNMGLSLAATTSPPOLLACGPTVHTCSENIVKGL 106  
Db 61 SLPCVXLDELFWNTSIENVTVEKNMTFGSTL-VTPNPGGLACGPLYAYRCGHLHYTGI 119

Qy 107 CFLFGNLAQQPQKPEALRGCPQEDSDIAFLDGSGLIIPDFRPMKEWSTVMEQLKK 166  
Db 120 CSDVSTPQVNSIAP-VQECSTQ-LDIVVLDSNSIYPWDS-----VTAFLNDLLK 170

Qy 167 -----SKTLFSLMYSSEFRHFTFKFQNNPNPSLIXPTQLLGR-THATGLRKVV 219  
Db 171 RMDIGPKQTVGIVQYGENVTHEFNILNKYSSTEEVLVAAKIVQRCGROTMTALGTDIAR 230

Qy 220 RELFNITNGARKNAFKILPILTDGKEFGPLGYEDVPELDREGVIRYVIGYDAFR--- 276  
Db 231 KEAFTARGARRGKVKWIVITDGESH-DNHRUKVIOQCEDNIQRFIALIGLSYNRGN 289

Qy 277 --SEKROELINTVASKPRPDHVPQINNFEALKTIONQLREKIPAIPTGTQSGSSSEHEM 334  
Db 290 LSTKEFVEEIKSTASSTPEKHFNVSDDELAVIIVKTGLGERIPALEATADQSAASPSEM 349

Qy 335 SQGFGAAITNSGPLISTVGSYDAGGVPLYTSKE-----KSTP-INMTFVDSMDMDAYL 388  
Db 350 SQTGFSAHSQDVMGLGAVGWNGTVVQKASQIIIPNTTFNVSTKKNPL-ASYL 408

Qy 389 GYAAAIILNRVQSL-VLGAAPRYQHILGNAMPQNTGMNESNANVGTQIGAYFGSGLCS 447  
Db 409 GYTVNSATASSGDVLIYACOPRYNHTGVIIYRMEDGNKILQTLSEGLIGSYFGSILTT 468

Qy 448 VDVSNGSTDLVLIGAPHY-----YEQTR-GGQVSVCPPLPRGORARWQCD 491  
Db 469 TDIDKDSNTDILLVGAPWYMTGTEKEBQGVVYVALNQTRFEYQVSLPIKQTCSSROHN 528

Qy 492 AVLYGEGQGPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLFGTSGSGISP 550  
Db 529 SCTTENKNEPCGARFGTRIAAAMDNLNDGFNDIVIGAPLEDDGGAVYIHG-SGTIRK 587

Qy 551 SHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRLSQPVLRYKAIM 610  
Db 588 EYAQRIPSGDGTLLKFPFGQSIHGMEDLNGDLTDTVTIGLGAALFWSRDVAVVKVTWN 647

Qy 511 PNPREVARNVPCNDQVVKKEAG--EVRVCLHVO-KSTEDRLREGQIQSVVTVOLALDS 667  
Db 648 FEPKNVNIQKNCH---MEGKETVCINATVCFEVLKSKEDTIYEADLQ-----YRVTLDS 700

Qy 668 GRPHSRVAFNET-----KNSTRQTOVLGELTQCTELKQLPNCIEDPVSPIVLRNPS 721  
Db 701 LRQISRSFFSGTQERKQYRNITVRKSEC-----TKHSFYMLDKHDQDSVR---ITLDFN 752

Qy 722 LVGTPLSAFNLRPVLADARLFTALFPEKQKCGNDNICODDLSTIFPSMSLDCLVVG 781  
Db 753 LT-DPENG-----PVLDDSLPNSVHEYIPFAKOCGNKKEKICISLSLHVATTEKDLIVRS 806

Qy 782 PRE-FNVTVTVRNDGESSYQTVTFPPPLDLVYKRVKSTLQNRQSQRSLRACASSSTEV 840  
Db 807 QNDKFNLSLTVKNTKDSAYNTRIVHYSPLNLSVSGIEAIQKD-----SCSEN----- 853

Qy 841 SGALKSTSCSINHPIFENSEVFNITFDVDSKASIGN-KLLKANVTSENMPRTNKE 899

854 ---HNITCKVGYFPLRRGEMVTKILFQFNTSYLMENVTIYLSATSDSEBPPETLSDNV 909  
Qy 900 FQLELPVKYAYVAVVTSCHGVSTKVLNFTASEVTSMVQHOYQVSN-----LGQRS----- 949  
Db 910 VNISSIPVKEVGLQFYS-SASEYHISIAANTVPEVINSTEDIGREINIFILIRKSGSP 968  
Qy 950 ---LPISLVP-----LVPVRLNQTVIWRPQVTFSENLSTCHTKE----- 987

Db 969 MPBLKLSISFPMNITSGVPLYPTGLSS-----SENANCRPHIFEDPPFSINGK 1017  
Qy 988 RLPSHSDFLABLRKAPVWNCIAVQCORIQCDIPFGIOE-----BFN 1029

Db 1018 KMTTSTD---HLKRGTTILDCTCKEATITCLTASDISQVNVSLILMKPTIKSYFSSLN 1074  
Qy 1030 ATLKNLSFDWVIKTSNHLIIVSTABILPNDVSTLLPGQCAFVRSTETKVFPEFVN 1089

Db 1075 LTRIGEL-----RSENASIVLSSN-----QKRELAIQISKDGLPG 1110

Qy 1090 PLPL--IVGSGVGLLILALITALYKLGFFKQYKMMSE 1128  
Db 1111 RVPLAVILLSAPAGILLMLLMLLILALWKIGFFKRPPLKKMEK 1151

RESULT 8

145914  
Integrin alpha 2 subunit - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Sep-2003  
C:Accession: I45914  
R:Kanata, T.; Puzon, W.; Takada, Y.  
J. Biol. Chem. 269, 9659-9663, 1994  
A:Title: Identification of putative ligand binding sites within the I-domain of integrin  
A:Reference number: A54402; PMID:94193647; PMID:7511592  
A:Accession: I45914  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1170 <KAM>  
A:Cross-references: GB:I45886; NID:9438695; PIDN:AAB59255.1; PID:9439696  
F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.2%; Score 1071; DB 2; Length 1170;  
Best Local Similarity 27.4%; Pred. NO. 2.4e-66;  
Matches 333; Conservative 219; Mismatches 495; Indels 169; Gaps 47;

Qy 1 FNLDTENAMTFQ-ENARGFGQSVQL---QGSRVVVGAPQEIIVAAQNRGLSQCC--DYST 54  
Db 19 YNVGLPKAKIFSGPSSSGFVAVQQFINPKGNWLLVGSFWSGFPKRRMGDVYKCPDVLST 78

Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLAATTSFPQLACGPTVHTCSENIVKGLC 107  
Db 79 TTCXLNLQTSMSNVTEMKTNMSLGLTLTRNVGTGGFLTCGPLMAQCCGQYVYTVGC 138

Qy 108 FLFGNLAQQPQKPEALRGCPQEDSDIAFLDGSGLIIPDFRPMKEWSTVMEQLK-- 165  
Db 139 SDVSPDF-QLRTSFAPAVQTCF-SPIDVVVVCDESNISYPWD--AVKAFLEKFXVQGLDIG 194

Qy 166 KSKTLFSLMYSSEFRHFTFKFQNNPNPSLIXPTQLL-----GRHTATGLRKVRNE 221  
Db 195 PRTQMGILLQYANNRVVFNLTNFKSKD---EMIKATSQTPQYGGDLTWTPEAIQVARDT 251

Qy 222 LFNITNGARKNAFKILFLTDGKEFGDPLGYEDVPELDREGVIRYVIGV-----GDAFR 276  
Db 252 AYSTAAGRPGKATKVVVVVVDGESH-DGSKLKAVIDQCNKNIIRFGLAVGLYLNALD 310

Qy 277 SEKRSQELNIVASKPRPDHVPFOINNFEALKTIONQLREKIPAIPTGTQSGSSSEHEMSQ 336  
Db 311 TKNLKEIKALIASITERHFPNVSDADLLEKAGTIGQIFSIGTVQCG-GNFMQWMSQ 369

Qy 337 EGFSAAIT--SNGPLLSTVGSYDAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGYA 391  
Db 370 VGFSAEYSQNNILMULGAVDWSGTVVQKTPHGLIFSKOAFEQILQDRHNSSYLGS 429



QY 956 FLV-----EVLNQTIVDRPOVTF-SENLS 980  
 DB 985 LVTHIPIQYTKENPLLYLGTGQDQAGDISCTAEINPLKLPHTA-----PSVSFRNENR 1040  
 QY 981 STCHTKRLPSHDFLAELKAPVWNGSIACVQRIQCDIPFGIQBEFNATLKNLSFDM 1040  
 DB 1041 ---HTKE-----LDCTTSCSNITCMLKDLHMQAEYFINVTRVNRRT 1080  
 QY 1041 YIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEKVPFPEVNPPLPIVGVSSVG 1100  
 DB 1081 PAASFTQVQLTAAAEIDTHNPQLFVIEENAVIPLIMKPTKAEVPT--GVILGIIIA 1138  
 QY 1101 GILLALITAAALYKLGFFKQYKDM 1125  
 DB 1139 GILLALLAMTAGLWLGFFKQYKDM 1163

## RESULT 10

A35854  
 integrin alpha-1 chain precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 23-Oct-1990 #sequence\_revision 13-Sep-1991 #text\_change 15-Sep-2003  
 C:Accession: A35854; S11243  
 R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 703-720, 1990  
 A:Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin A  
 A:Reference number: A35854; MUID:90339125; PMID:2380249  
 A:Accession: A35854  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1180 <IGN>  
 A:Cross-references: GB:X52140; NID:956493; PIDN:CAA36384.1; PID:956494  
 C:Keywords: cell adhesion; cytoskeleton; transmembrane protein  
 P:170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 17.9%; Score 1055; DB 2; Length 1180;  
 Best Local Similarity 27.4%; Pred. No. 3.2e-65;  
 Matches 346; Conservative 196; Mismatches 475; Indels 246; Gaps 48;  
 QY 1 FNLDTENAMTPOENARG-FQGVVQL---QGSRVVVGAPQEIIVAAQMGSLQCYDSTGS 56  
 DB 29 FNVDSKMSFSGPVEDMGFTVQYVENEGKWLIGSLVPGQKARTGDVYKCPVGRER 88  
 QY 57 CEP-IRLOVPEA-----VNKSLGLSLAATTSPQALLAGPTVHOTCSNTYVKGL 106  
 DB 89 AMPCVKLDLPVNTSIPNVTIENMTFGSTL-VTNPGGFLACGPLYAYRCGHLHYTGI 147  
 QY 107 CFLRGLMLRQOPKFPALAGCPQEDSDIAFLIDGGSIIPHDFRMKEWNVSTM---EQ 163  
 DB 148 CSDVSPFTQVNGFAP--VQECSTQ-LDIVIVLDGNSIYP-----WESVIAFLNDL 196  
 QY 164 LKK-----SKTLFSLMYSSEFRIHFTKFEQNPAPRSILKIPITQLLG-RHTTANGLRK 217  
 DB 197 LKRMIDGPKQTQGVIGYGVNTHEPNLANKYSSTEEVLAANKIGRQGLQTMALGIDT 256  
 QY 218 VVRELFTNGARKNAFKILPLTLDGKFGDPLGIEDVPELDREGVIRVYVGVGDAPR- 276  
 DB 257 ARKEAFTARGAGVKKVMVIVTDGESH-DNYELKQVIOCDCEDENIQFSTAILGHYNR 315  
 QY 277 ----SEKSRQLNVAASKPRDVFQINNFPAKTKTONQLREKIPALEOTQTGSSSPFH 332  
 DB 316 GNLSTEKFEVEIKSIASEPTEKFFNVSDALAVTVKALGERIFALEATADQSAASFEM 375  
 QY 333 EMSQEGFSAAITSNGPLLSVGVSDWAGGVFLITSKEKSTFFINMT--RVDSDMND---AY 387  
 DB 376 EMSQTGFSAHYSQDMVLMGAVGAYDMNGTVVMQKANKQVIPHNTTQTETPAKNGEPLASY 435  
 QY 388 LGYAAAILNRVQSLVGLAPRYOHTGLVAMFRQNTGMMESNANVKGTOIGAYFGASLCS 447  
 DB 436 LGYTVNSATIPGDVLYTAGOPRYNHTGQVVIYRMEDGNINILQTAGGEIGSYFGSVLTT 495  
 QY 448 VDVDSNGSTDLVLGAPHY-----YEQTR-GGQVSVCPPLPQORARWQCD 491

DB 496 IDIDKDSYTDLLLLGAPMYMGTKEKQKVVVYVAVNQTRPEYQMSLEPIRQTCSSSLKON 555  
 QY 492 AVLVEGOQOPWG-RFGAALTVLGVNVDGDKLTDVAIGAPEGEDNRGAVLYLFGTSSGISP 550  
 DB 556 SCTKENKNEPCGAFGTAIAAVKDLNVDGPNVDVIGAPLEDDEHAGAVYIYHG-SGKTIRE 614  
 QY 551 SHSORIAGSKLSPLOVFGOSLSGGODLTWDGVLVDLTVGAGQGHVLLRLRSOPVLRVKRAIME 610  
 DB 615 AVAQRIPSGGDKTLKFFGQSIHGEDMNGDGLTDTVTIGLGAALFWARDVAVVKVTWN 674  
 QY 611 FNPREVARNVPCNDQVVKKEAG--EVRVCLHVQ-KSTDRDLREGQIQSVTYDLDLDS 667  
 DB 675 FEPKNVNIQKNCR---VEGKTVTCINATWCFHVKLKSKEDSYEADLQ---YRVTLDS 727  
 QY 668 GRPHSRVAFNET-----KNSTRRTQVGLGLTQTCETKLQLPNCI-----EDPVS 712  
 DB 728 LRQISRSPFGSQTKERKIQRNITVRESE-----CIHSPWMLDKHDFOD 770  
 QY 713 PIVLRINPESLVGTPLSAPGNLRVULADAQRLFTALPPPEKNGNDNI CODDLSITFSEM 772  
 DB 771 SVRVTLDFNLT-DENG-----PVLDDALPNSVHEHIPPAKDCNKKERCSLDUTLNVSTT 824  
 QY 773 SLDCLVVGGPRE-PNVTVTVVRNDGEDSYRTQVTFPPFPDLSYRKVSTLQMRQSRWRLA 831  
 DB 825 EKSLLVKSQHDKNFVSLTVKNGDSAYNTRTVVQHSNLI FSGIEBIQKD-----S 876  
 QY 832 CESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLL-LKANVTSEN 890  
 DB 877 CESN-----QNITCRVGYPLRAGETVTKIIPQFNTPSHLSENAILHLSATSDSEE 927  
 QY 891 NMPRTNTEFOLELPKYAV---YMWVTSHGVSST-----KYLNFASSENTSRVMOHQ 939  
 DB 928 PLESNDNEVNISSIPVRYEVLQFYSSASBHHISVAANETIPEFINST--EDIGNEINVF 985  
 QY 940 YQVSNLQORSIP---ISLVP-----IVPRLNQTWIND-----RP----- 971  
 DB 986 YTRKRGRHFWPELQLSISFPNLTADGYPLVIG-----WSSSNVNCRPSLEDDPFG 1039  
 QY 972 ----QVTFPS-----ENLSSTCHTKERLPSHSDFLAELKAPVWNCISAVCQRIQ 1016  
 DB 1040 INSGKMTISKSEVLKRGITDCSSTC-----GVATITCSLLPSDLSQ 1082  
 QY 1017 CDI-----PPFGIQEFP---NATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFTLL 1067  
 DB 1083 VVSLALLWKPTF-IRAHFSSIALTLRGELX-----SENSITLSSN----- 1123  
 QY 1068 PQGAFVRSQTEKVPPEVNPPLPL--IVGSSVGGILLIALITAAALYKLGFFKQYKDM 1125  
 DB 1124 -----RKRELAQISKDGLPGRVPLWVILLSAPAGLMLLMLLALWKIGFFKRLPKK 1177  
 QY 1126 MSE 1128  
 DB 1178 MEK 1180

## RESULT 11

A33998  
 integrin alpha-2 chain precursor - human  
 N:Alternate names: CD49b; platelet glycoprotein GPIIb; VLA-2/collagen receptor alpha-2 ct  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Mar-1990 #sequence\_revision 18-Sep-1992 #text\_change 15-Sep-2003  
 C:Accession: A33998; B56793; A53117  
 R:Takada, Y.; Hemler, M.E.  
 J. Cell Biol. 109, 397-407, 1989  
 A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet  
 A:Reference number: A33998; MUID:89306879; PMID:2545729  
 A:Accession: A33998  
 A:Molecule type: mRNA  
 A:Residues: 1-1181 <PAK>  
 A:Cross-references: GB:X17033; NID:933906; PIDN:CAA34894.1; PID:933907  
 A>Note: The authors translated the codon GAT for residue 802 as Gln, GTC for residue 803  
 R:Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.  
 Biochem. J. 279, 419-425, 1991



```
QY 96 TCSNTYVVKGLCLPFGSNLRQOPKQPFKPEALRGCPQEDSDIAFLIDGSGIIPHDFRMKE 155
DB 161 MKSDNKLPTGICVMPDRLTSLK
QY 156 WTSVWMEQLKSKTLFSLQYSEEPRIHFTFKFQNNPNRSLIKPITQLLGRTHATCL 215
DB 186 -----NMFCYKDYT-----
QY 216 RKVRELFNITNGARKNAFKILFLTDGKFGDPLGYDVIPELDREGVIRYVIGVDAP 275
DB 196 -----RKFG-----
QY 276 RSEKSRQELNIVASKPRDHVPQINNFPEALKTIONQLREKIFAIEGTQTGSSSFEHMS 335
DB 201 -----NFAS-----
QY 336 QGFSAAITNSGPLSTVGSYDWAGVFLY--TSKEKSTFINNTRVDSMDNDAYLGYA-- 391
DB 206 QAGISSFYTDLLVMGAPGSSYWTGVVYVNTTQYKAFVD--RQNVKFGSLGYSGV 263
QY 392 AAILLRNVSQVLGAPRYOHGLVAMFRONTGMESNANV-----KGTGIGAYFGASLCS 447
DB 264 AGHFRSPHTTEVVGAPQEOHQKAYIF-----SIDENELNIYEMKKGKLSYFGASVCA 319
QY 448 VDVDSNGSTDVLIGAPHYVEQTRGQVSVCPPLRGORA--RWQCDVAVLGEQCPWGRFG 506
DB 320 VLNADGFSDL-LVGAPMSTIREGRVFVY--INSGGAVWEMERVLVGSQKYA--ARFG 376
QY 507 AALTVDGVNGDKLTDVAGAPHEENRGAVYLFHGTSGSGISPHSQRISAKLSPLRIQ 566
DB 377 ESTANGLDINDGFDFAIGAPQEDDLRGAVIYNGRV--DGISSTYSQRIEQQISKSLR 435
QY 567 YGQSLSGGQDLTMCGLDLVTGCA--QGHVLLRSGPLRVKALMEFNPVARNVFCN 624
DB 436 MFGQSSGGQIDANNNGYVDVAGAFQSDSAVLLRTPVIVVEASLS--HPEVNRKFDCT 494
QY 625 DQVVGKEAGEVRVCLHVQKSTRDLREQIQISVVTYDIALDSGR---PHSRAVF--NET 679
DB 495 -----ENGLPSVCVHLLCFYSYKKEVPGVIVLFPVYVSLDVRHKAESPFSFYFSGT 547
QY 680 KNSRTRQVGLGTQCETLKLQFNCIEDPVSPIVLRNPLS-----VGTPLSAPGMLR 734
DB 548 SDVITGSIRVSSGKCRTHQAFMRKQVDRDILTPIHVEATYHLGHVHTKRNTEFFPLQ 607
QY 735 PVLAEDAQR-LFTALPFPKNCNDNICQDLSITPS-----FMSLDCLVVGGPREFNV 787
DB 608 PILQOKKEQVIRKMINPAPFCAVEN--CSADLQVSAKVGLFKEYNTYLAVGSKTIML 666
QY 788 TVTVRNDGDSYRQVTFPPFLDLSYRKVSTIQNQRSQSWRLACESASTVSGALKST 847
DB 667 NVSLFNAGDDAYETTLNVQLPTGLYPIKILDLLEK-----QINCE---VTSSGIVK-L 716
QY 848 SCISINHPFPENSEVTFNITEDVDSKASIGNKLLKAVTSEN--NMPRTNTEFQLELP 905
DB 717 ACSLGHVYDRLSRDISFLDVSLSRAHEDLSISVHASCENEGELDQVRNRTVLIP 776
QY 906 VYAVTVMYTSHGV--STKYLNFTASENTSRYMQH-----YQVNLGQSLP--LSLVEL 957
DB 777 LAYEV--MLTVHGLVNPSTFVYGSSENEPECTMAEKLNLTFHVTINTGISMAPNVSVKIM 834
QY 958 VVVRNQTIVMDRP--QVTFSENLSSTCHKE-----RLPSSHSDFLAELR 1000
DB 835 VP-----NSFLPQDDKLFNLDVQTTTQCCHFKHVGRECTFAQKGAGTUDIVLKSLKD 891
QY 1001 KAPVNVNCIA--VCQRIQCDIPFGIQEBFNATLKGNLSPDWIYIKTSHNHLIVASTABIL 1058
DB 892 KR--LLYCMKADQHCDFLCN-----FGWMSG-----KEASVHIQLEGRPSIL 933
QY 1059 FNDSVTLAPGQAFVRSQTEHTKVBEPVNP-----
DB 934 EMDETSSL-----KFEIKATAPPEPHKPVIELNKDENVAVHFLGLHHRQPKRH 983
```

```
QY 1091 -LPIIVGSSVGLLALITAAALYKLGFFKROYKDMSE 1128
DB 984 TIIITISLLGLLVLLISCVMMKAGFFKROYKSILQE 1022

RESULT 13
S06046
Integrin alpha-4 chain precursor - human
N:Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C:Species: Homo sapiens (man)
C:Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 24-Sep-1999
C:Accession: S06046; A39355; D28018
R:Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A:Reference number: S06046; MUID:89356603; PMID:2788572
A:Accession: S06046
A:Molecule type: mRNA
A:Residues: 1-1038 <TAK>
A:Cross-references: GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1; PID:g33946
R:Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A:Title: Characterization of the alpha-4 integrin gene promoter.
A:Reference number: A39355; MUID:91239513; PMID:2034655
A:Accession: A39355
A:Molecule type: DNA
A:Residues: 1-93 <ROS>
A:Cross-references: GB:M62841
R:Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A:Title: The very late antigen family of heterodimers is part of a superfamily of molec
A:Reference number: A94151; MUID:87204112; PMID:3033641
A:Accession: D28018
A:Molecule type: protein
A:Residues: 40-50, 'E', 52-53 <TA2>
C:Genetics:
A:Gene: GDB:ITGA4; CD49D
A:Cross-references: GDB:128032; OMIM:192975
A:Map position: 2q31-2q32
C:Superfamily: integrin alpha-4 chain
C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transm
F:1-39/Domain: signal sequence #status predicted <SIG>
F:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.9%; Score 642; DB 2; Length 1038;
Best Local Similarity 23.8%; Pred. No. 2.3e-36;
Matches 293; Conservative 175; Mismatches 409; Indels 356; Gaps 53;

QY 1 ENLDTENATFQ--ENARFGQSVV--QLQGSR--VVVGAPOEIVAN-----QSGSLYQCDY 52
DB 40 YNVDTESALLYGGPHNTLFGYSVVHLHSGANRLLVGAPTANLANASVINPGAIYRCRI 99
QY 53 STG---SCEPIRLQVPVBAVNMVSLGLSLAATTSPPQLLACGFTVHQTCSENTYVKGCLCF 109
DB 100 GKNPCQTCQQLQIGSP-----NGBP-----KTCLEBRDNCWL--- 133
QY 110 FGSNLRQOPKFPFALRGCPQEDSDIAFLIDGSGIIPHDFRMKEWSTVWMEQLKSKT 169
DB 134 -GVTLRQPGS-----NGSIVTCGHR-----W-----KN 156
QY 170 LFSLMQYSEEPRIHFTFKFQNNPNRSLIKPITQLLGRTHATCLRVKVRLEFNITNGA 229
DB 157 IFYI-----
QY 230 RKNAPKILFLTDEKFGDPLGYSDVTPPELDREGVIRYVIGVDAPFRSEKSRQELNIVAS 289
DB 161 -KNEKNK---LPTGCGYGP-----PDL-----RTELSEKRIA 187
QY 230 KPPRDHVPQI--NNPEALKTIONQLREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGP 348
DB 188 PCYQDVVKFGFENFAS-----COAGISSFTYKDLI 217
QY 349 LLSTVGSYDWAGGVFLY---TSKEKSTFINNTRVDSMDNDAYLGYA---AAILLRNVSQ 403
```



Db 918 1Y-----MLANT-BILKDDSSVIO-----FMSRAKVKYDPALRVVBEIAHGNPEEV 962  
Qy 1085 ---FEVPR---PLPLIVG-----SSVGGLLLLALITAAALYKLGFFKROYKDM 1126  
Db 963 TWTFEALHLEPRGVVGVGWIATISLILVGLIFLLAVLLVWKGFFRRYKEII 1015  
RESULT 15  
T31437  
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)  
C:Species: Lytechinus variegatus (variegated urchin)  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
C/Accession: T31437  
R:Hertzler, P.L.; McClay, D.R.  
submitted to the EMBL Data Library, May 1998  
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.  
A:Reference number: Z21035  
A:Accession: T31437  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1041 <HER>  
A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC23572.1  
A:Experimental source: developmental stage embryo  
C:Function:  
A:Description: binds laminin  
C:Superfamily: integrin alpha-2b chain  
Query Match 10.4%; Score 614.5; DB 2; Length 1041;  
Best Local Similarity 24.8%; Pred. No. 2e-34;  
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;  
Qy 336 QEGPSAALTNGP--LLSTGVSYDWAGGVFLVTSKEKTEFNMTVRDS-----DMNDAYLG 389  
Db 131 QAGSGIIFSDNSALVMAGPSYVLOGQIYVQSLLNRSV-VQATQESNTGYSDNSYRG 239  
Qy 330 YAAAI--ILNRVQSLVIGAPRYOHI-GLVAFRQNTGMNESNANVKTOIGAYFGASLC 446  
Db 240 YSLALGDFNGDGVQDVVVGTPRAESLMGLVAIFDQNLNQPFI---QVMGTQIVAYFGYSVT 296  
Qy 447 SVDVDSNGSTDLVLIGAPHYEQTREGGOVSVCPLPRGQARMCQDVALYGEQ----- 498  
Db 297 VVDI--NNDYDOLLVGAAPYMDGPAIQ-----RWEAGAVVYVQLNPDVGPGA 343  
Qy 499 -----GQWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGS 546  
Db 344 SNRLSLSTLIGGOIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDDAGAVYVHG-SAN 402  
Qy 547 GISPSHSQRIAGSLIS-PRLOVFGQSLSGQDLTMDGLVDTVGAQ--CHVLLILRSQVYL 603  
Db 403 GLKSTPAQVLTPTSLGHSGITTFGFSLQGGQMDKNKYFDLLVGAESANTAVLIRTPWV 462  
Qy 604 RVKALMEFNPREVARNVPECNDQVYKGEAGEVRVCLHVOKSTRDRREGQIQSV----- 658  
Db 463 SLDATLNTPE-----IGINLENKTYE-LADGTMTVTSFIAMT 497  
Qy 659 -----VTYDLALDSG-RPHSRVFNTRKNTSRQTVLGL-TQTCETLKLQ 702  
Db 498 CFTYTGNYLPDHIDISYTVTVDSGLIANRRAMFVNDNMSEITKTRRLAVSTQFCDPLEAY 557  
Qy 703 LPNCEIDFVSPVLRNLNLSLNGTPLSAFGN-----LRPLVAEDAQRLLTALFPPEX 753  
Db 558 VGNSTEDKLTPLKVTQYDL-----NDESRLOPHEIPLIDMATMSTQTKQVSIQN 609  
Qy 754 NCGNDNICQDLSITFSFMSLDCLVVGPRFNFVTVVRNDGEDSYRTQVTFPPPLDLSY 813  
Db 610 NCVN-NICIPDLDTVT-PNLPNI VIGQTQELTLDVSLNNGEDAFQSSLSVYVPLGLQF 667  
Qy 814 KXVSTLQNRQSRWELACASASSTEVSGALKSTSCSINHPIFFEN-----SEVTENIT 867  
Db 668 VRL-----ERKANMDFSVTCSDSD-----LRITCDTGNPMVGKNILEFGLTILSTFQVS 717  
Qy 868 FDVDS-----KASLGKLLKLLKANVTSENMPRTNKTEFQLELPVKYAVYVMTVSHGVSTK 922

Db 718 GKDSIEFYFKAESSENS--EDPNTLNNELNMTVPVTVDCITLKLKLSASYPEIVMYSTQED 775  
Qy 923 YL-----NFTASENTSRVWQHQYQV-----SNLQORSLPIS-----LVFLVPV 960  
Db 776 YVPPPPAKNASEADIGKEVM-HLYEVRNTGSSNAGEVSLNIQWPQKNEDGEVLYLLGI 834  
Qy 961 RLNQTVIWRPQ-----VTFSENLSTCTHKERLPSHSDFLAELRKAPV 1005  
Db 835 MTEBEGVTCQLTQGANPEGVKLEPSTKAKLSNSTTVSGRKRREPEVAEALAQTDN--VI 892  
Qy 1006 NCSIAVCORIQCQDIPFPFGIOBEFNAT-----LKGNLSPDWYIKTSHNLLIVSTABILF 1059  
Db 893 YCASDSQVLIINCTI-----DEINASKSVVRILGRF--W-----BRTF 928  
Qy 1060 NDSVFTLLPQGGAPVRSGTETKVE--PFEVNP-----LP----- 1092  
Db 929 QKAVSELTPVQVQATIASASAAVKTIPYNIPLPRDFSDSTKASTLVTEELVPPVTPIAW 988  
Qy 1093 -LIVGSSVGGLLLLALITAAALYKLGPFKR-----QYKDMMS--EGGPP 1132  
Db 989 WIIVSVLGGIILLILLILGLNKGPFERKKPGEKEKAYAVASADKGGPP 1038

Search completed: June 7, 2004, 17:18:02  
Job time : 38.559 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:55:38 ; Search time 9.97194 Seconds  
(without alignments)  
5937.039 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5894

Sequence: 1 FNLDTENATFQENARGFGQ.....FKRQYKDNMSGGPPGASPPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5836.5	99.2	1152	1 ITAM HUMAN	P11215 homo sapien
2	4460	75.8	1153	1 ITAM MOUSE	P05555 mus musculus
3	3459	58.8	1163	1 ITAX HUMAN	P20702 homo sapien
4	3401	57.8	1162	1 ITAD HUMAN	Q13349 homo sapien
5	1547.5	26.3	1170	1 ITAL HUMAN	P20701 homo sapien
6	1527.5	26.0	1163	1 ITAL MOUSE	P24063 mus musculus
7	1148.5	19.5	1167	1 ITAE MOUSE	Q60677 mus musculus
8	1140	19.4	1179	1 ITAE HUMAN	P38570 homo sapien
9	1093.5	18.6	1151	1 ITAI HUMAN	P56199 homo sapien
10	1084.5	18.4	1189	1 ITAH HUMAN	Q9UXK5 homo sapien
11	1071	18.2	1170	1 ITAZ BOVIN	P53710 bos taurus
12	1057	18.0	1178	1 ITA2 MOUSE	Q62469 mus musculus
13	1055	17.9	1180	1 ITAI RAT	P18614 rattus norv
14	1054	17.9	1181	1 ITA2 HUMAN	P17301 homo sapien
15	1051.5	17.9	1167	1 ITAG HUMAN	Q75578 homo sapien
16	663	11.3	1039	1 ITA4 MOUSE	Q00651 mus musculus
17	642	10.9	1038	1 ITA4 HUMAN	P13512 homo sapien
18	630	10.7	1035	1 ITA4 HUMAN	Q13797 homo sapien
19	593.5	10.1	1032	1 ITA4 XENLA	Q91687 xenopus lae
20	571.5	9.7	1066	1 ITA3 CRISP	P17852 cricetidae
21	567.5	9.6	1053	1 ITA3 MOUSE	Q62470 mus musculus
22	555.5	9.4	1053	1 ITA5 MOUSE	P11688 mus musculus
23	546.5	9.3	1034	1 ITAV CHICK	P26008 gallus gall
24	545.5	9.3	1050	1 ITA5 XENLA	Q06274 xenopus lae
25	537.5	9.1	1130	1 ITA6 HUMAN	P23429 homo sapien
26	535	9.1	1044	1 ITAV MOUSE	P43406 mus musculus
27	534.5	9.1	1072	1 ITA6 CHICK	P26007 gallus gall
28	532.5	9.0	1066	1 ITA3 HUMAN	P26006 homo sapien
29	532	9.0	1049	1 ITA5 HUMAN	P08648 homo sapien
30	526	8.9	1048	1 ITAV HUMAN	P06756 homo sapien
31	517	8.8	1044	1 ITA8 CHICK	P26009 gallus gall
32	512.5	8.7	1091	1 ITA6 MOUSE	P61739 mus musculus
33	498	8.5	1179	1 ITA7 MOUSE	Q61738 mus musculus

#### ALIGNMENTS

##### RESULT 1

ID	ITAM HUMAN	STANDARD	PRT	1152 AA.
AC	P11215			
DT	01-JUN-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor Mo1)			
DE	(Neutrophil adherence receptor)			
GN	ITGAM OR CR3A OR CD11B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=883115033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Renold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.P., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RP	MEDLINE=89098993; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,			
RT	Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).			
RN	[6]			

34	494	8.4	1396	1 ITA2 DROME	P12080 drosophila
35	491.5	8.4	1146	1 ITAL DROME	Q24247 drosophila
36	490	8.3	1033	1 ITAB MOUSE	Q9QUM0 mus musculus
37	489.5	8.3	1039	1 ITAB HUMAN	P08514 homo sapien
38	489	8.3	1025	1 ITAB HUMAN	P53708 homo sapien
39	486	8.3	126	1 ITAM CAVEO	P11578 cavia porce
40	471	8.0	1181	1 ITA7 HUMAN	Q13683 homo sapien
41	470	8.0	1106	1 ITA7 RAT	Q63258 rattus norv
42	462	7.9	1226	1 PAT2 CAEEL	P34446 caenorhabdi
43	444.5	7.6	1139	1 INAL CAEEL	Q03600 caenorhabdi
44	427	7.3	1115	1 ITA3 DROME	O4386 drosophila
45	385	6.5	1000	1 ITA5 DROME	Q9W1M8 drosophila

RP SEQUENCE OF 1-9 FROM N.A.  
RX MEDLINE=92073318; PubMed=1683702;  
RA Shelley C.S., Arnaut M.A.;  
RT "The promoter of the CD11b gene directs myeloid-specific and  
RT developmentally regulated expression.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).  
RN [7]  
RP SEQUENCE OF 1-9 FROM N.A.  
RX TISSUE=Blood;  
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;  
RT "Characterization of the myeloid-specific CD11b promoter.";  
RL Blood 79:865-870(1992).  
RN [8]  
RP SEQUENCE OF 17-31.  
RX MEDLINE=87076671; PubMed=3539202;  
RA Pierce M.W., Remold-O'Donnell E., Todd R.P. III, Arnaut M.A.;  
RT "N-terminal sequence of human leukocyte glycoprotein Mol.  
RT conservation across species and homology to platelet IIb/IIIa.";  
RL Biochim. Biophys. Acta 874:368-371(1986).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.  
RX MEDLINE=95171458; PubMed=7867070;  
RA Lee J.O., Rieu P., Arnaut M.A., Liddington R.;  
RT "Crystal structure of the A domain from the alpha subunit of integrin  
RT CR3 (CD11b/CD18).";  
RL Cell 80:631-638(1995).  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.  
RX MEDLINE=96363671; PubMed=8747460;  
RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;  
RT "Two conformations of the integrin A-domain (I-domain): a pathway for  
RT activation?";  
RL Structure 3:1333-1340(1995).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.  
RX MEDLINE=98362595; PubMed=9687375;  
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,  
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,  
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,  
RA Muchler V.I., Tomach C.S., Watenpaugh K.D., Wiley V.H.;  
RT "Cation binding to the integrin CD11b I domain and activation model  
RT assessment.";  
RL Structure 6:923-935(1998).  
RN [12]  
RP 3D-STRUCTURE MODELING OF 17-616.  
RX MEDLINE=9826734; PubMed=9560195;  
RA Oxvig C., Springer T.A.;  
RT "Experimental support for a beta-propeller domain in integrin alpha-  
RT subunits and a calcium binding site on its lower surface.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).  
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS  
CC ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES  
CC AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.  
CC IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF  
CC THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D  
CC PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR  
CC FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES  
CC OF FIBRINOGEN GAMMA CHAIN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M  
CC ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND  
CC GRANULOCYTES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- DATABASE: NEMO-PRO; NOTE-CD guide CD11b entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: J03925; AAA59544.1; -  
CC EMBL: M18044; AAA59491.1; -  
CC EMBL: J04145; AAA59903.1; -  
CC EMBL: S52227; AAB24821.1; -  
CC EMBL: S52152; AAB24821.1; JOINED.  
CC EMBL: S52153; AAB24821.1; JOINED.  
CC EMBL: S52154; AAB24821.1; JOINED.  
CC EMBL: S52155; AAB24821.1; JOINED.  
CC EMBL: S52157; AAB24821.1; JOINED.  
CC EMBL: S52159; AAB24821.1; JOINED.  
CC EMBL: S52161; AAB24821.1; JOINED.  
CC EMBL: S52164; AAB24821.1; JOINED.  
CC EMBL: S52165; AAB24821.1; JOINED.  
CC EMBL: S52167; AAB24821.1; JOINED.  
CC EMBL: S52169; AAB24821.1; JOINED.  
CC EMBL: S52170; AAB24821.1; JOINED.  
CC EMBL: S52173; AAB24821.1; JOINED.  
CC EMBL: S52174; AAB24821.1; JOINED.  
CC EMBL: S52180; AAB24821.1; JOINED.  
CC EMBL: S52181; AAB24821.1; JOINED.  
CC EMBL: S52184; AAB24821.1; JOINED.  
CC EMBL: S52189; AAB24821.1; JOINED.  
CC EMBL: S52191; AAB24821.1; JOINED.  
CC EMBL: S52192; AAB24821.1; JOINED.  
CC EMBL: S52203; AAB24821.1; JOINED.  
CC EMBL: S52212; AAB24821.1; JOINED.  
CC EMBL: S52213; AAB24821.1; JOINED.  
CC EMBL: S52216; AAB24821.1; JOINED.  
CC EMBL: S52219; AAB24821.1; JOINED.  
CC EMBL: S52220; AAB24821.1; JOINED.  
CC EMBL: S52221; AAB24821.1; JOINED.  
CC EMBL: S52222; AAB24821.1; JOINED.  
CC EMBL: S52226; AAB24821.1; JOINED.  
CC EMBL: M76724; AAA58410.1; -  
CC EMBL: M84477; AAA51960.1; -  
CC PIR: A31108; RWHU1B.  
CC PDB: 1A8X; 17-JUN-98.  
CC PDB: 1BRO; 18-NOV-98.  
CC PDB: 1BHQ; 18-NOV-98.  
CC PDB: 1IDN; 25-NOV-98.  
CC PDB: 1IDO; 01-AUG-96.  
CC PDB: 1JLM; 11-JAN-97.  
CC PDB: 1MIU; 07-AUG-02.  
CC Genew: HGNC:6149; ITGAM.  
CC MIM: 120980; -  
CC GO: GO:0008305; C: integrin complex; TAS.  
CC GO: GO:0004895; F: cell adhesion receptor activity; TAS.  
CC GO: GO:0007155; P: cell adhesion; TAS.  
CC InterPro: IPR000413; Integrin\_alpha.  
CC InterPro: IPR002035; VWFA.  
CC Pfam: PF01839; FG-GAP; 3.  
CC Pfam: PF00357; Integrin\_A; 1.  
CC Pfam: PF00092; vwfa; 1.  
CC PRINTS: PRO1185; INTEGRINA.  
CC PRINTS: PRO0453; VWFADOMAIN.  
CC SMART: SM00191; Int\_alpha; 4.  
CC SMART: SM00327; VWFA; 1.  
CC PROSITE: PS00242; INTEGRIN\_ALPHA; 1.  
CC PROSITE: PS50234; VWFA; 1.  
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
CC Signal; 3D-structure; Repeat; Magnesium; Calcium.  
CC SIGNAL 1 16  
CC CHAIN 17 1152 INTEGRIN ALPHA-M.  
CC DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 1105 1128 POTENTIAL.  
CC DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).  
CC

Query Match		99.2%; Score 5836.5; DB 1; Length 1152;	
Best Local Similarity		99.1%; Pred. No. 0;	
Matches 1127; Conservative		7; Mismatches	2; Indels 1; Gaps 1;
QY	1	FNLDTENAMTFQENARGFQSQVVLQGSVVVVGAGQEIIVAAHQRSLSVOCYSTGSCPEI	60
DB	17	FNLDTENAMTFQENARGFQSQVVLQGSVVVVGAGQEIIVAAHQRSLSVOCYSTGSCPEI	76
QY	61	RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFGLNLRQOQK	120
DB	77	RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFGLNLRQOQK	136
QY	121	FPBALRGCCQEDSDIAFLIDGSGSIIPDFRMEKWSVTMEOLKSKTKSLFSLMOYSEF	180
DB	137	FPBALRGCCQEDSDIAFLIDGSGSIIPDFRMEKWSVTMEOLKSKTKSLFSLMOYSEF	196
QY	181	RIEFTPFKEFQNNPNRSLKIPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFILL	240
DB	197	RIEFTPFKEFQNNPNRSLKIPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFILL	256
QY	241	TGEXFGDPLGYEDVPELDREGVIRYVTVGSDAFRSEKSRQELATVASKPPRDHVFQIN	300
DB	257	TGEXFGDPLGYEDVPELDREGVIRYVTVGSDAFRSEKSRQELATVASKPPRDHVFQIN	316
QY	301	NFEALKTIONQLREKIFAETGTQTSSTSFHEMSQESFSAITNGPLLSVTGSDWAG	360
DB	317	NFEALKTIONQLREKIFAETGTQTSSTSFHEMSQESFSAITNGPLLSVTGSDWAG	376
QY	361	GVFLTSKSKSTFINNTRVDSMDNDAYLGAAAILNRVQSLVLGAPRYQHIGLVAMFR	420
DB	377	GVFLTSKSKSTFINNTRVDSMDNDAYLGAAAILNRVQSLVLGAPRYQHIGLVAMFR	436
QY	421	QNTGWMESNANVKGTGIGAYFGASLCSVDVDNGSTDLVLGAPHYETQRCQSVCP	480
DB	437	QNTGWMESNANVKGTGIGAYFGASLCSVDVDNGSTDLVLGAPHYETQRCQSVCP	496
QY	481	PRQGRARWCDVLYGEQOQPMGRGAALTVLGDVNGDKLTDVAICAPEDENRGAVYLF	540
DB	497	PRG-BARWQCDVLYGEQOQPMGRGAALTVLGDVNGDKLTDVAICAPEDENRGAVYLF	555
QY	541	HGTSGSGISPSHSQRIAGSKLPRLOYFGOSLGGQDLTMDGLVDLTWGAQGHVLLRSQ	600
DB	556	HGTSGSGISPSHSQRIAGSKLPRLOYFGOSLGGQDLTMDGLVDLTWGAQGHVLLRSQ	615
QY	601	PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
DB	616	PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	675
QY	661	YDLALDSGPHSRVAFNETKSTRQTQVLGTLQTCETLKLQLPNCIEDPVSPIVRLNF	720
DB	676	YDLALDSGPHSRVAFNETKSTRQTQVLGTLQTCETLKLQLPNCIEDPVSPIVRLNF	735
QY	721	SLVGTPLSAFGNLRPVLAEDAQLPTALPPFKNCNDNI CODDLSITFSFMSLDCLVVG	780
DB	736	SLVGTPLSAFGNLRPVLAEDAQLPTALPPFKNCNDNI CODDLSITFSFMSLDCLVVG	795
QY	781	GPREFNVTVVNDGEDSVRTQVTFPPFLDLRYKVTSLQNRQSRWLACESASSTEV	840
DB	796	GPREFNVTVVNDGEDSVRTQVTFPPFLDLRYKVTSLQNRQSRWLACESASSTEV	855
QY	841	SGALKSTCSINHPIFPENSEVTENITPDVDSKASLGKLLKANVTSENMPRTNKTEP	900
DB	856	SGALKSTCSINHPIFPENSEVTENITPDVDSKASLGKLLKANVTSENMPRTNKTEP	915
QY	901	QLELPVKAVVMVTSKYLNTASNTSRVWQHQQYQVSNLQSRSLPSLFLVPLV	960
DB	916	QLELPVKAVVMVTSKYLNTASNTSRVWQHQQYQVSNLQSRSLPSLFLVPLV	975
QY	961	RLNQTVMIDRPQVTSNLSTCTKRLPSHSDFLAELRKAPVNCIAVCORIQCDIP	1020
DB	976	RLNQTVMIDRPQVTSNLSTCTKRLPSHSDFLAELRKAPVNCIAVCORIQCDIP	1035

1021 PFGIOEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSOTET 1080

1036 PFGIOEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSOTET 1095

1081 KVEPEFVNPPLIVGSSVGLLALITAAALYKLGFKKQYKDKMSGGGPGGAEPO 1137

1096 KVEPEFVNPPLIVGSSVGLLALITAAALYKLGFKKQYKDKMSGGGPGGAEPO 1152

RESULT 2

ITAM MOUSE STANDARD; PRT; 1153 AA.

AC P0555;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).

GN ITGAM.

OS Mus musculus (Mouse);

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN SEQUENCE FROM N.A.

RP MEDLINE=88312584; PubMed=3044779;

RA Pytela R.;

RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";

RL EMBO J. 7:1371-1378(1988).

RN [1]

RP SEQUENCE OF 11-45 FROM N.A.

RC STRAIN=BALB/c; TISSUE=Spleen;

EX MEDLINE=85287312; PubMed=2942940;

RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.W., Springer T.A.;

RT "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";

RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).

RN [2]

RP SEQUENCE OF 17-28.

EX MEDLINE=8518276; PubMed=3887182;

RA Springer T.A., Teplow D.B., Dreyer W.J.;

RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon.";

RL Nature 314:540-542(1985).

CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN MAST CELL DEVELOPMENT AND IN IMMUNE COMPLEX-MEDIATED GLOMERULONEPHRITIS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-M SUBUNIT GENE DEMONSTRATE INCREASE IN NEUTROPHIL ACCUMULATION, IN RESPONSE TO AN IMPAIRED DEGRADATION AND PHAGOCYTOSIS, EVENTS THAT APPARENTLY ACCELERATE APOPTOSIS IN NEUTROPHILS. THESE MICE DEVELOP OBESITY.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN. ALPHA-M ASSOCIATES WITH BETA-2.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND GRANULOCYTES.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 1 VMFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; X07640; CAA30479.1; -  
 CC EMBL; M14293; AAA39484.1; -  
 CC PIR; S00551; S00551.  
 CC HSP; P11215; IABX.  
 CC MGSP; MGI:196607; Itgam.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP; 3.  
 CC Pfam; PF00357; Integrin\_A; 1.  
 CC Pfam; PF00092; vwa; 1.  
 CC PRINTS; PR01185; INTEGRINA.  
 CC PRINTS; PR00453; VWFADOMAIN.  
 CC SMART; SM00191; Int\_alpha; 5.  
 CC SMART; SM00327; VWA; 1.  
 CC PROSITE; PS00242; INTEGRIN ALPHA; 1.  
 CC PROSITE; PS00234; VWA; 1.  
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 CC Signal; Calcium; Repeat.  
 CC SIGNAL 1  
 CC CHAIN 17 1153  
 CC DOMAIN 17 1105  
 CC TRANSMEM 1106 1129  
 CC DOMAIN 1130 1153  
 CC REPEAT 31 84  
 CC REPEAT ? ?  
 CC DOMAIN 164 350  
 CC REPEAT 337 400  
 CC REPEAT 401 452  
 CC REPEAT 454 515  
 CC REPEAT 517 575  
 CC REPEAT 580 632  
 CC REPEAT 645 673  
 CC CA\_BIND 529 537  
 CC CA\_BIND 592 600  
 CC SITE 1132 1136  
 CC DISULFID 66 73  
 CC DISULFID 105 123  
 CC DISULFID 654 711  
 CC DISULFID 770 776  
 CC DISULFID 999 1023  
 CC DISULFID 1028 1033  
 CC CARBOHYD 58 58  
 CC CARBOHYD 86 86  
 CC CARBOHYD 391 391  
 CC CARBOHYD 696 696  
 CC CARBOHYD 734 734  
 CC CARBOHYD 772 772  
 CC CARBOHYD 801 801  
 CC CARBOHYD 881 881  
 CC CARBOHYD 907 907  
 CC CARBOHYD 941 941  
 CC CARBOHYD 980 980  
 CC CARBOHYD 994 994  
 CC CARBOHYD 1022 1022  
 CC CARBOHYD 1045 1045  
 CC CARBOHYD 1051 1051  
 CC CARBOHYD 1076 1076  
 CC SEQUENCE 1153 AA; 127480 MW; 178DB988ABCEB0343 CRC64;  
 Query Match 75.88; Score 4460; DB 1; Length 1153;  
 Best Local Similarity 73.88; Pred. No. 1.8e-290;  
 Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;

CC 1 FNLDTEHPMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 60  
 CC 17 FNLDTEHPMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 76

QY 61 RLOVPEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGICPLGSGNLQQPOK 120  
 DB 77 PIQVPEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGICPLGSGNLQQPOK 136  
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEMWSTVMBQKSKTKLPGLMQLYSEF 180  
 DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEMWSTVMBQKSKTKLPGLMQLYSEF 196  
 QY 181 RIHFTFK3QNNPNRSLIKPITOLLGRTHATGKRVKVRRELFNITNGARKNAFKILPLL 240  
 DB 197 RIHFTFNDPKRNPSPRSHVSPKIQNGRTKTASGIRKVVRELFNITNGARKNAFKILPLL 256  
 QY 241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDGAFRSKSKQELATVASKPRDHFVQIN 300  
 DB 257 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDGAFRSKSKQELATVASKPRDHFVQIN 316  
 QY 301 NFEALKTIONQLREKIPALEGTQTCSSSFHEMSQEGFSAITNSGPLSTVSGVDNAG 360  
 DB 317 NFEALKTIONQLREKIPALEGTQTCSSSFHEMSQEGFSAITNSGPLSTVSGVDNAG 376  
 QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGZAAAIILNRVQSLVGLGAPRYCHIGLVAMFR 420  
 DB 377 GAFLYTSKDKVTFINNTRVDSMDNDAYLGZAAAIILNRVQSLVGLGAPRYCHIGLVAMFR 436  
 QY 421 QNTGAWESNANVKGTOIGAYFGASLCSDVDNSGSTDILVIGAPHYVETRGGOVSCPL 480  
 DB 437 ENFGTWBPHSTSIKSGQIGSYFGASLCSDVDNSGSTDILVIGAPHYVETRGGOVSCPL 496  
 QY 481 PRGQARWOCDAVLYGBOGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEEENRGAVLPL 540  
 DB 497 PRGQARWOCDAVLYGBOGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEEENRGAVLPL 555  
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTVGAQCHVLLRSQ 600  
 DB 556 YGASTASLSASHSHRIIGAHFSPGLQYFGQSLGSGQDLTMDGLVDLTVGAQCHVLLRSQ 615  
 QY 601 PVLRYKALMEFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
 DB 616 PVLRYKALMEFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675  
 QY 661 YDLALDSGRPHSRVAFNETKSTRTQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 720  
 DB 676 YDLALDPVRSIRAPFDETKNTTRTQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 735  
 QY 721 SLVGTPLSAFGLNLRVLAEDAQRFTALFPPEKNGNDNICODDLSITFSEWSDCLVVG 780  
 DB 736 TLVGTPLSAFGLNLRVLAEDAQRFTALFPPEKNGNDNICODDLSITFSEWSDCLVVG 795  
 QY 781 GPREFNVTVVRNDGEDSYRTQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 839  
 DB 796 GPREFNVTVVRNDGEDSYRTQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 855  
 QY 840 VSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKT 899  
 DB 856 VSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKT 915  
 QY 900 FQLELPVXYAVVWVTSRGVSTKYLNFTASENTSRVMOHQVQVSNLQORSLPISLVELVP 959  
 DB 916 FQLELPVXYAVVWVTSRGVSTKYLNFTASENTSRVMOHQVQVSNLQORSLPISLVELVP 975  
 QY 960 VRLNQTIVWDRPQVTPSENLSSTCHTKERLPSHSDFLAELRKA PVVNCISIAVCRIQCDI 1019  
 DB 976 VRLNQTIVWDRPQVTPSENLSSTCHTKERLPSHSDFLAELRKA PVVNCISIAVCRIQCDI 1035  
 QY 1020 PFFGIQEBFENATLKNLSFDWYIKTSHNELITVSTAEILFVDSVFTLPLPGQAFVRSQTE 1079  
 DB 1036 PFFGIQEBFENATLKNLSFDWYIKTSHNELITVSTAEILFVDSVFTLPLPGQAFVRSQTE 1095  
 QY 1080 TKVEPFEVNPVPLIVGSGVGLLLALITAALYKLGFFKQYKDMGEGGPPGAPQ 1137  
 DB 1096 TKVEPFEVNPVPLIVGSGVGLLLALITAALYKLGFFKQYKDMGEGGPPGAPQ 1153

RESULT 3  
 ITEX\_HUMAN STANDARD; PRT; 1163 AA.  
 AC P20702;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95  
 DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).  
 GN ITEX OR CD11c.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=88166645; PubMed=3327687;  
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;  
 RT "cDNA cloning and complete primary structure of the alpha subunit of  
 RT a leukocyte adhesion glycoprotein, p150,95";  
 RL EMBO J. 6:4023-4028(1987).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=90153906; PubMed=2303426;  
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;  
 RT "Genomic structure of an integrin alpha subunit, the leukocyte  
 RT p150,95 molecule";  
 RL J. Biol. Chem. 265:2782-2788(1990).  
 RN [3]  
 RN ERRATUM  
 RP Corbi A.L., Garcia-Aguilar J., Springer T.A.;  
 RA J. Biol. Chem. 265:12750-12751(1990).  
 RN [4]  
 RN SEQUENCE OF 20-43.  
 RX MEDLINE=87167596; PubMed=3549901;  
 RA Miller L.J., Wiebe M., Springer T.A.;  
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1  
 RT and p150,95 leukocyte adhesion proteins";  
 RL J. Immunol. 138:2381-2383(1987).  
 CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT  
 CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL  
 CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY  
 CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X  
 CC ASSOCIATES WITH BETA-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND  
 CC GRANULOCYTES.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M81695; AA59180.1; -;  
 DR EMBL; Y00093; CA66283.1; -;  
 DR EMBL; M29165; -; NOT ANNOTATED CDS.  
 DR EMBL; M29487; AA51620.1; ALT SEQ.  
 DR EMBL; M29482; AA51620.1; JOINED.  
 DR EMBL; M29483; AA51620.1; JOINED.  
 DR EMBL; M29484; AA51620.1; JOINED.  
 DR EMBL; M29485; AA51620.1; JOINED.

DR EMBL; M29486; AA51620.1; JOINED.  
 DR EMBL; M29487; AA51620.1; JOINED.  
 DR EMBL; M29488; AA51620.1; JOINED.  
 DR EMBL; M29489; AA51620.1; JOINED.  
 DR EMBL; M29490; AA51620.1; JOINED.  
 DR EMBL; M29491; AA51620.1; JOINED.  
 DR EMBL; M29492; AA51620.1; JOINED.  
 DR EMBL; M29493; AA51620.1; JOINED.  
 DR EMBL; M29494; AA51620.1; JOINED.  
 DR EMBL; M29495; AA51620.1; JOINED.  
 DR EMBL; M29496; AA51620.1; JOINED.  
 DR EMBL; M29497; AA51620.1; JOINED.  
 DR EMBL; M29498; AA51620.1; JOINED.  
 DR EMBL; M29499; AA51620.1; JOINED.  
 DR EMBL; M29500; AA51620.1; JOINED.  
 DR EMBL; M29501; AA51620.1; JOINED.  
 DR EMBL; M29502; AA51620.1; JOINED.  
 DR EMBL; M29503; AA51620.1; JOINED.  
 DR EMBL; M29504; AA51620.1; JOINED.  
 DR EMBL; M29505; AA51620.1; JOINED.  
 DR EMBL; M29506; AA51620.1; JOINED.  
 DR EMBL; M29507; AA51620.1; JOINED.  
 DR EMBL; M29508; AA51620.1; JOINED.  
 DR EMBL; M29509; AA51620.1; JOINED.  
 DR EMBL; M29510; AA51620.1; JOINED.  
 DR EMBL; M29511; AA51620.1; JOINED.  
 DR EMBL; M29512; AA51620.1; JOINED.  
 DR EMBL; M29513; AA51620.1; JOINED.  
 DR EMBL; M29514; AA51620.1; JOINED.  
 DR EMBL; M29515; AA51620.1; JOINED.  
 DR EMBL; M29516; AA51620.1; JOINED.  
 DR EMBL; M29517; AA51620.1; JOINED.  
 DR EMBL; M29518; AA51620.1; JOINED.  
 DR EMBL; M29519; AA51620.1; JOINED.  
 DR EMBL; M29520; AA51620.1; JOINED.  
 DR EMBL; M29521; AA51620.1; JOINED.  
 DR EMBL; M29522; AA51620.1; JOINED.  
 DR EMBL; M29523; AA51620.1; JOINED.  
 DR EMBL; M29524; AA51620.1; JOINED.  
 DR EMBL; M29525; AA51620.1; JOINED.  
 DR EMBL; M29526; AA51620.1; JOINED.  
 DR EMBL; M29527; AA51620.1; JOINED.  
 DR EMBL; M29528; AA51620.1; JOINED.  
 DR EMBL; M29529; AA51620.1; JOINED.  
 DR EMBL; M29530; AA51620.1; JOINED.  
 DR EMBL; M29531; AA51620.1; JOINED.  
 DR EMBL; M29532; AA51620.1; JOINED.  
 DR EMBL; M29533; AA51620.1; JOINED.  
 DR EMBL; M29534; AA51620.1; JOINED.  
 DR EMBL; M29535; AA51620.1; JOINED.  
 DR EMBL; M29536; AA51620.1; JOINED.  
 DR EMBL; M29537; AA51620.1; JOINED.  
 DR EMBL; M29538; AA51620.1; JOINED.  
 DR EMBL; M29539; AA51620.1; JOINED.  
 DR EMBL; M29540; AA51620.1; JOINED.  
 DR EMBL; M29541; AA51620.1; JOINED.  
 DR EMBL; M29542; AA51620.1; JOINED.  
 DR EMBL; M29543; AA51620.1; JOINED.  
 DR EMBL; M29544; AA51620.1; JOINED.  
 DR EMBL; M29545; AA51620.1; JOINED.  
 DR EMBL; M29546; AA51620.1; JOINED.  
 DR EMBL; M29547; AA51620.1; JOINED.  
 DR EMBL; M29548; AA51620.1; JOINED.  
 DR EMBL; M29549; AA51620.1; JOINED.  
 DR EMBL; M29550; AA51620.1; JOINED.  
 DR EMBL; M29551; AA51620.1; JOINED.  
 DR EMBL; M29552; AA51620.1; JOINED.  
 DR EMBL; M29553; AA51620.1; JOINED.  
 DR EMBL; M29554; AA51620.1; JOINED.  
 DR EMBL; M29555; AA51620.1; JOINED.  
 DR EMBL; M29556; AA51620.1; JOINED.  
 DR EMBL; M29557; AA51620.1; JOINED.  
 DR EMBL; M29558; AA51620.1; JOINED.  
 DR EMBL; M29559; AA51620.1; JOINED.  
 DR EMBL; M29560; AA51620.1; JOINED.  
 DR EMBL; M29561; AA51620.1; JOINED.  
 DR EMBL; M29562; AA51620.1; JOINED.  
 DR EMBL; M29563; AA51620.1; JOINED.  
 DR EMBL; M29564; AA51620.1; JOINED.  
 DR EMBL; M29565; AA51620.1; JOINED.  
 DR EMBL; M29566; AA51620.1; JOINED.  
 DR EMBL; M29567; AA51620.1; JOINED.  
 DR EMBL; M29568; AA51620.1; JOINED.  
 DR EMBL; M29569; AA51620.1; JOINED.  
 DR EMBL; M29570; AA51620.1; JOINED.  
 DR EMBL; M29571; AA51620.1; JOINED.  
 DR EMBL; M29572; AA51620.1; JOINED.  
 DR EMBL; M29573; AA51620.1; JOINED.  
 DR EMBL; M29574; AA51620.1; JOINED.  
 DR EMBL; M29575; AA51620.1; JOINED.  
 DR EMBL; M29576; AA51620.1; JOINED.  
 DR EMBL; M29577; AA51620.1; JOINED.  
 DR EMBL; M29578; AA51620.1; JOINED.  
 DR EMBL; M29579; AA51620.1; JOINED.  
 DR EMBL; M29580; AA51620.1; JOINED.  
 DR EMBL; M29581; AA51620.1; JOINED.  
 DR EMBL; M29582; AA51620.1; JOINED.  
 DR EMBL; M29583; AA51620.1; JOINED.  
 DR EMBL; M29584; AA51620.1; JOINED.  
 DR EMBL; M29585; AA51620.1; JOINED.  
 DR EMBL; M29586; AA51620.1; JOINED.  
 DR EMBL; M29587; AA51620.1; JOINED.  
 DR EMBL; M29588; AA51620.1; JOINED.  
 DR EMBL; M29589; AA51620.1; JOINED.  
 DR EMBL; M29590; AA51620.1; JOINED.  
 DR EMBL; M29591; AA51620.1; JOINED.  
 DR EMBL; M29592; AA51620.1; JOINED.  
 DR EMBL; M29593; AA51620.1; JOINED.  
 DR EMBL; M29594; AA51620.1; JOINED.  
 DR EMBL; M29595; AA51620.1; JOINED.  
 DR EMBL; M29596; AA51620.1; JOINED.  
 DR EMBL; M29597; AA51620.1; JOINED.  
 DR EMBL; M29598; AA51620.1; JOINED.  
 DR EMBL; M29599; AA51620.1; JOINED.  
 DR EMBL; M29600; AA51620.1; JOINED.  
 DR EMBL; M29601; AA51620.1; JOINED.  
 DR EMBL; M29602; AA51620.1; JOINED.  
 DR EMBL; M29603; AA51620.1; JOINED.  
 DR EMBL; M29604; AA51620.1; JOINED.  
 DR EMBL; M29605; AA51620.1; JOINED.  
 DR EMBL; M29606; AA51620.1; JOINED.  
 DR EMBL; M29607; AA51620.1; JOINED.  
 DR EMBL; M29608; AA51620.1; JOINED.  
 DR EMBL; M29609; AA51620.1; JOINED.  
 DR EMBL; M29610; AA51620.1; JOINED.  
 DR EMBL; M29611; AA51620.1; JOINED.  
 DR EMBL; M29612; AA51620.1; JOINED.  
 DR EMBL; M29613; AA51620.1; JOINED.  
 DR EMBL; M29614; AA51620.1; JOINED.  
 DR EMBL; M29615; AA51620.1; JOINED.  
 DR EMBL; M29616; AA51620.1; JOINED.  
 DR EMBL; M29617; AA51620.1; JOINED.  
 DR EMBL; M29618; AA51620.1; JOINED.  
 DR EMBL; M29619; AA51620.1; JOINED.  
 DR EMBL; M29620; AA51620.1; JOINED.  
 DR EMBL; M29621; AA51620.1; JOINED.  
 DR EMBL; M29622; AA51620.1; JOINED.  
 DR EMBL; M29623; AA51620.1; JOINED.  
 DR EMBL; M29624; AA51620.1; JOINED.  
 DR EMBL; M29625; AA51620.1; JOINED.  
 DR EMBL; M29626; AA51620.1; JOINED.  
 DR EMBL; M29627; AA51620.1; JOINED.  
 DR EMBL; M29628; AA51620.1; JOINED.  
 DR EMBL; M29629; AA51620.1; JOINED.  
 DR EMBL; M29630; AA51620.1; JOINED.  
 DR EMBL; M29631; AA51620.1; JOINED.  
 DR EMBL; M29632; AA51620.1; JOINED.  
 DR EMBL; M29633; AA51620.1; JOINED.  
 DR EMBL; M29634; AA51620.1; JOINED.  
 DR EMBL; M29635; AA51620.1; JOINED.  
 DR EMBL; M29636; AA51620.1; JOINED.  
 DR EMBL; M29637; AA51620.1; JOINED.  
 DR EMBL; M29638; AA51620.1; JOINED.  
 DR EMBL; M29639; AA51620.1; JOINED.  
 DR EMBL; M29640; AA51620.1; JOINED.  
 DR EMBL; M29641; AA51620.1; JOINED.  
 DR EMBL; M29642; AA51620.1; JOINED.  
 DR EMBL; M29643; AA51620.1; JOINED.  
 DR EMBL; M29644; AA51620.1; JOINED.  
 DR EMBL; M29645; AA51620.1; JOINED.  
 DR EMBL; M29646; AA51620.1; JOINED.  
 DR EMBL; M29647; AA51620.1; JOINED.  
 DR EMBL; M29648; AA51620.1; JOINED.  
 DR EMBL; M29649; AA51620.1; JOINED.  
 DR EMBL; M29650; AA51620.1; JOINED.  
 DR EMBL; M29651; AA51620.1; JOINED.  
 DR EMBL; M29652; AA51620.1; JOINED.  
 DR EMBL; M29653; AA51620.1; JOINED.  
 DR EMBL; M29654; AA51620.1; JOINED.  
 DR EMBL; M29655; AA51620.1; JOINED.  
 DR EMBL; M29656; AA51620.1; JOINED.  
 DR EMBL; M29657; AA51620.1; JOINED.  
 DR EMBL; M29658; AA51620.1; JOINED.  
 DR EMBL; M29659; AA51620.1; JOINED.  
 DR EMBL; M29660; AA51620.1; JOINED.  
 DR EMBL; M29661; AA51620.1; JOINED.  
 DR EMBL; M29662; AA51620.1; JOINED.  
 DR EMBL; M29663; AA51620.1; JOINED.  
 DR EMBL; M29664; AA51620.1; JOINED.  
 DR EMBL; M29665; AA51620.1; JOINED.  
 DR EMBL; M29666; AA51620.1; JOINED.  
 DR EMBL; M29667; AA51620.1; JOINED.  
 DR EMBL; M29668; AA51620.1; JOINED.  
 DR EMBL; M29669; AA51620.1; JOINED.  
 DR EMBL; M29670; AA51620.1; JOINED.  
 DR EMBL; M29671; AA51620.1; JOINED.  
 DR EMBL; M29672; AA51620.1; JOINED.  
 DR EMBL; M29673; AA51620.1; JOINED.  
 DR EMBL; M29674; AA51620.1; JOINED.  
 DR EMBL; M29675; AA51620.1; JOINED.  
 DR EMBL; M29676; AA51620.1; JOINED.  
 DR EMBL; M29677; AA51620.1; JOINED.  
 DR EMBL; M29678; AA51620.1; JOINED.  
 DR EMBL; M29679; AA51620.1; JOINED.  
 DR EMBL; M29680; AA51620.1; JOINED.  
 DR EMBL; M29681; AA51620.1; JOINED.  
 DR EMBL; M29682; AA51620.1; JOINED.  
 DR EMBL; M29683; AA51620.1; JOINED.  
 DR EMBL; M29684; AA51620.1; JOINED.  
 DR EMBL; M29685; AA51620.1; JOINED.  
 DR EMBL; M29686; AA51620.1; JOINED.  
 DR EMBL; M29687; AA51620.1; JOINED.  
 DR EMBL; M29688; AA51620.1; JOINED.  
 DR EMBL; M29689; AA51620.1; JOINED.  
 DR EMBL; M29690; AA51620.1; JOINED.  
 DR EMBL; M29691; AA51620.1; JOINED.  
 DR EMBL; M29692; AA51620.1; JOINED.  
 DR EMBL; M29693; AA51620.1; JOINED.  
 DR EMBL; M29694; AA51620.1; JOINED.  
 DR EMBL; M29695; AA51620.1; JOINED.  
 DR EMBL; M29696; AA51620.1; JOINED.  
 DR EMBL; M29697; AA51620.1; JOINED.  
 DR EMBL; M29698; AA51620.1; JOINED.  
 DR EMBL; M29699; AA51620.1; JOINED.  
 DR EMBL; M29700; AA51620.1; JOINED.  
 DR EMBL; M29701; AA51620.1; JOINED.  
 DR EMBL; M29702; AA51620.1; JOINED.  
 DR EMBL; M29703; AA51620.1; JOINED.  
 DR EMBL; M29704; AA51620.1; JOINED.  
 DR EMBL; M29705; AA51620.1; JOINED.  
 DR EMBL; M29706; AA51620.1; JOINED.  
 DR EMBL; M29707; AA51620.1; JOINED.  
 DR EMBL; M29708; AA51620.1; JOINED.  
 DR EMBL; M29709; AA51620.1; JOINED.  
 DR EMBL; M29710; AA51620.1; JOINED.  
 DR EMBL; M29711; AA51620.1; JOINED.  
 DR EMBL; M29712; AA51620.1; JOINED.  
 DR EMBL; M29713; AA51620.1; JOINED.  
 DR EMBL; M29714; AA51620.1; JOINED.  
 DR EMBL; M29715; AA51620.1; JOINED.  
 DR EMBL; M29716; AA51620.1; JOINED.  
 DR EMBL; M29717; AA51620.1; JOINED.  
 DR EMBL; M29718; AA51620.1; JOINED.  
 DR EMBL; M29719; AA51620.1; JOINED.  
 DR EMBL; M29720; AA51620.1; JOINED.  
 DR EMBL; M29721; AA51620.1; JOINED.  
 DR EMBL; M29722; AA51620.1; JOINED.  
 DR EMBL; M29723; AA51620.1; JOINED.  
 DR EMBL; M29724; AA51620.1; JOINED.  
 DR EMBL; M29725; AA51620.1; JOINED.  
 DR EMBL; M29726; AA51620.1; JOINED.  
 DR EMBL; M29727; AA51620.1; JOINED.  
 DR EMBL; M29728; AA51620.1; JOINED.  
 DR EMBL; M29729; AA51620.1; JOINED.  
 DR EMBL; M29730; AA51620.1; JOINED.  
 DR EMBL; M29731; AA51620.1; JOINED.  
 DR EMBL; M29732; AA51620.1; JOINED.  
 DR EMBL; M29733; AA51620.1; JOINED.  
 DR EMBL; M29734; AA51620.1; JOINED.  
 DR EMBL; M29735; AA51620.1; JOINED.  
 DR EMBL; M29736; AA51620.1; JOINED.  
 DR EMBL; M29737; AA51620.1; JOINED.  
 DR EMBL; M29738; AA51620.1; JOINED.  
 DR EMBL; M29739; AA51620.1; JOINED.  
 DR EMBL; M29740; AA51620.1; JOINED.  
 DR EMBL; M29741; AA51620.1; JOINED.  
 DR EMBL; M29742; AA51620.1; JOINED.  
 DR EMBL; M29743; AA51620.1; JOINED.  
 DR EMBL; M29744; AA51620.1; JOINED.  
 DR EMBL; M29745; AA51620.1; JOINED.  
 DR EMBL; M29746; AA51620.1; JOINED.  
 DR EMBL; M29747; AA51620.1; JOINED.  
 DR EMBL; M29748; AA51620.1; JOINED.  
 DR EMBL; M29749; AA51620.1; JOINED.  
 DR EMBL; M29750; AA51620.1; JOINED.  
 DR EMBL; M29751; AA51620.1; JOINED.  
 DR EMBL; M29752; AA51620.1; JOINED.  
 DR EMBL; M29753; AA51620.1; JOINED.  
 DR EMBL; M29754; AA51620.1; JOINED.  
 DR EMBL; M29755; AA51620.1; JOINED.  
 DR EMBL; M29756; AA51620.1; JOINED.  
 DR EMBL; M29757; AA51620.1; JOINED.  
 DR EMBL; M29758; AA51620.1; JOINED.  
 DR EMBL; M29759; AA51620.1; JOINED.  
 DR EMBL; M29760; AA51620.1; JOINED.  
 DR EMBL; M29761; AA51620.1; JOINED.  
 DR EMBL; M29762; AA51620.1; JOINED.  
 DR EMBL; M29763; AA51620.1; JOINED.  
 DR EMBL; M29764; AA51620.1; JOINED.  
 DR EMBL; M29765; AA51620.1; JOINED.  
 DR EMBL; M29766; AA51620.1; JOINED.  
 DR EMBL; M29767; AA51620.1; JOINED.  
 DR EMBL; M29768; AA51620.1; JOINED.  
 DR EMBL; M29769; AA51620.1; JOINED.  
 DR EMBL; M29770; AA51620.1; JOINED.  
 DR EMBL; M29771; AA51620.1; JOINED.  
 DR EMBL; M29772; AA51620.1; JOINED.  
 DR EMBL; M29773; AA51620.1; JOINED.  
 DR EMBL; M29774; AA51620.1; JOINED.  
 DR EMBL; M29775; AA51620.1; JOINED.  
 DR EMBL; M29776; AA51620.1; JOINED.  
 DR EMBL; M29777; AA51620.1; JOINED.  
 DR EMBL; M29778; AA51620.1; JOINED.  
 DR EMBL; M29779; AA51620.1; JOINED.  
 DR EMBL; M29780; AA51620.1; JOINED.  
 DR EMBL; M29781; AA51620.1; JOINED.  
 DR EMBL; M29782; AA51620.1; JOINED.  
 DR EMBL; M29783; AA51620.1; JOINED.  
 DR EMBL; M29784; AA51620.1; JOINED.  
 DR EMBL; M29785; AA51620.1; JOINED.  
 DR EMBL; M29786; AA51620.1; JOINED.  
 DR EMBL; M29787; AA51620.1; JOINED.  
 DR EMBL; M29788; AA51620.1; JOINED.  
 DR EMBL; M29789; AA51620.1; JOINED.  
 DR EMBL; M29790; AA51620.1; JOINED.  
 DR EMBL; M29791; AA51620.1; JOINED.  
 DR EMBL; M29792; AA51620.1; JOINED.  
 DR EMBL; M29793; AA51620.1; JOINED.  
 DR EMBL; M29794; AA51620.1; JOINED.  
 DR EMBL; M29795; AA51620.1; JOINED.  
 DR EMBL; M29796; AA51620.1; JOINED.  
 DR EMBL; M29797; AA51620.1; JOINED.  
 DR EMBL; M29798; AA51620.1; JOINED.  
 DR EMBL; M29799; AA51620.1; JOINED.  
 DR EMBL; M29800; AA51620.1; JOINED.  
 DR EMBL; M29801; AA51620.1; JOINED.  
 DR EMBL; M29802; AA51620.1; JOINED.  
 DR EMBL; M29803; AA51620.1; JOINED.  
 DR EMBL; M29804; AA51620.1; JOINED.  
 DR EMBL; M29805; AA51620.1; JOINED.  
 DR EMBL; M29806; AA51620.1; JOINED.  
 DR EMBL; M29807; AA51620.1; JOINED.  
 DR EMBL; M29808; AA51620.1; JOINED.  
 DR EMBL; M29809; AA51620.1; JOINED.  
 DR EMBL; M29810; AA51620.1; JOINED.  
 DR EMBL; M29811; AA51620.1; JOINED.  
 DR EMBL; M29812; AA51620.1; JOINED.  
 DR EMBL; M29813; AA51620.1; JOINED.  
 DR EMBL; M29814; AA51620.1; JOINED.  
 DR EMBL; M29815; AA51620.1; JOINED.  
 DR EMBL; M29816; AA51620.1; JOINED.  
 DR EMBL; M29817; AA51620.1; JOINED.  
 DR EMBL; M29818; AA51620.1; JOINED.  
 DR EMBL; M29819; AA51620.1; JOINED.  
 DR EMBL; M29820; AA51620.1; JOINED.  
 DR EMBL; M29821; AA51620.1; JOINED.  
 DR EMBL; M29822; AA51620.1; JOINED.  
 DR EMBL; M29823; AA51620.1; JOINED.  
 DR EMBL; M29824; AA51620.1; JOINED.  
 DR EMBL; M29825; AA51620.1; JOINED.  
 DR EMBL; M29826; AA51620.1; JOINED.  
 DR EMBL; M29827; AA51620.1; JOINED.  
 DR EMBL; M29828; AA51620.1; JOINED.  
 DR EMBL; M29829; AA51620.1; JOINED.  
 DR EMBL; M29830; AA51620.1; JOINED.  
 DR EMBL; M29831; AA51620.1; JOINED.  
 DR EMBL; M29832; AA51620.1; JOINED.  
 DR EMBL; M29833; AA51620.1; JOINED.  
 DR EMBL; M29834; AA51620.1; JOINED.  
 DR EMBL; M29835; AA51620.1; JOINED.  
 DR EMBL; M29836; AA51620.1; JOINED.  
 DR EMBL; M29837; AA51620.1; JOINED.  
 DR EMBL; M29838; AA51620.1; JOINED.  
 DR EMBL; M29839; AA51620.1; JOINED.  
 DR EMBL; M29840; AA51620.1; JOINED.  
 DR EMBL; M29841; AA51620.1; JOINED.  
 DR EMBL; M29842; AA51620.1; JOINED.  
 DR EMBL; M29843; AA51620.1; JOINED.  
 DR EMBL; M29844; AA51620.1; JOINED.  
 DR EMBL; M29845; AA51620.1; JOINED.  
 DR EMBL; M29846; AA51620.1; JOINED.  
 DR EMBL; M29847; AA51620.1; JOINED.  
 DR EMBL; M29848; AA51620.1; JOINED.  
 DR EMBL; M29849; AA51620.1; JOINED.  
 DR EMBL; M29850; AA51620.1; JOINED.  
 DR EMBL; M29851; AA51620.1; JOINED.  
 DR EMBL; M29852; AA51620.1; JOINED.  
 DR EMBL; M29853; AA51620.1; JOINED.  
 DR EMBL; M29854; AA51620.1; JOINED.  
 DR EMBL; M29855; AA51620.1; JOINED.  
 DR EMBL; M29856; AA51620.1; JOINED.  
 DR EMBL; M29857; AA51620.1; JOINED.  
 DR EMBL; M29858; AA51620.1; JOINED.  
 DR EMBL; M29859; AA51620.1; JOINED.  
 DR EMBL; M29860; AA51620.1; JOINED.  
 DR EMBL; M29861; AA51620.1; JOINED.  
 DR EMBL; M29862; AA51620.1; JOINED.  
 DR EMBL; M29863; AA51620.1; JOINED.  
 DR EMBL; M29864; AA51620.1; JOINED.  
 DR EMBL; M29865; AA51620.1; JOINED.  
 DR EMBL; M29866; AA51620.1

Db 138 LPVSROECPQEQDIYFLIDGSGSISRNFATMNFVRAVISQFQSPSTQFSLMQFSNKE 197  
Qy 181 RIHFTKEFONNPNRSLIKPIITOLIGRTHATGLRKVVRELENIYNGARNAFKILFL 240  
Db 198 QTHFTFEFEERTNPLSLASVHLQLOQFTYTAIGNVVRLEFHASYGARRDAIKILIVI 257  
Qy 241 TDGKFGDPLGYEDVPELDREGVIRVIGVDAFRSEKROELNTVASKPRPDHVFQIN 300  
Db 258 TDGKKGDSLDYKDVIPADAAAGIIRYAGVGLAFQENSWKELNDIAKPSQEHIFKVE 317  
Qy 301 NFALAKTIQNLREKIFAIGTOTGSSSPHEHMSQEGFSAATISNGPLLSVTGSDVWAG 360  
Db 318 DFDALKIQNLREKIFAIGTETITSSSPHEHMSQEGFSAVTFPDPGVLGASFTWSG 377  
Qy 361 GVLTYTSKESKSTINTRVDSNDAYGLVAAAILRNVRQSLVGLCAPRYOHLGLVAFER 420  
Db 378 GAFLYPPNMSPTINNSQENVMRDSYLGSTELALWGVQSLVGLAPRYOHTGKAVIT 437  
Qy 421 QNTGMESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCEL 480  
Db 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEOTRGQVSVCEL 497  
Qy 481 PRGORARQCDVLYGEGQGPWCFGAALTVLGDVNDKLTQVAIGAPGEDNRGAVILF 540  
Db 498 PRGWR-RWQCDVLYGEGQGPWCFGAALTVLGDVNDKLTQVAIGAPGEDNRGAVILF 556  
Qy 541 HGTSGGSISSHSQRIAGSKLSPRLQYFGQSLGGQDLMGDLVDTYGAQGHVLLRSQ 600  
Db 557 HGVLPSPISPSHSQRIAGSKLSPRLQYFGQSLGGQDLMGDLVDTYGAQGHVLLRSQ 616  
Qy 601 PVLRYKALMEFNPREVARNVPECNDVQVKGAGEVRVCLVOKSSTDLRLEGQIOSVT 660  
Db 617 PVLWVGVMQFIPABIPRPAFECREQVQVSEQTIVQSNICLYIDKRSKLLGSDLOQSV 676  
Qy 661 YDLALDSRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEBPVSVILRLNF 720  
Db 677 LDALDPRGLSPRATQETKNSLSRVRLGLAKHCENFNLLPSCVEDSVYTIITLRLNF 736  
Qy 721 SIVGTPLSARGLNRPVLAEDAORLFRALPPEKQNGNDNICQDDLSITFSFMSLCLVVG 780  
Db 737 TLVCKPLAFNLRMLAALAAQRYFTASLPFEKNCAGADHICQDNLGISFSPGLKSLVG 796  
Qy 781 GPREFNVTVVRNDGDSYRQVTFPPFLDLSYRKVSTLQNSQSRSLRACSSASSTEV 840  
Db 797 SNLELNAEVMWNGEDSYCTITFSPHAGLSRYVAEGKQQLASLHLTCDSPVGV-- 854  
Qy 841 SGALKSTSCSINPIFENSEVTNITFDVDSKASLGNKLLKANTVSENMPRTKTEF 900  
Db 855 SQGTWSTSCRINHLIFRGAQITPLATFDVSPKAVLGDRLLLTANYSSENNTPRTSKTF 914  
Qy 901 QLELPVKYAVVMVTSHGVTCKYLNFTAS-ENTSRVMQHOYQVSNLQORSLPISLVFLVP 959  
Db 915 QLELPVKYAVYTVVSSEHQTKYLNFSSEKESHVAMHRYQNNLQORDLPVSNFWFP 974  
Qy 960 VRLNQTVINDRPOVTFSENLSSTCHTKERLPSHDSFLAELRKAPVWNCSTAVCQICDI 1019  
Db 975 VELNQAVMMDVEVSHPQNPFLRCSSEKIAPPASDFLAHTQKNPFLDCSIAGCLRFCDV 1034  
Qy 1020 PFCIOEFPNATKGNLSFMYIKTSHNLLIYSTAILFNDSVFTLLPQOGAFVSQTE 1079  
Db 1035 PFSVQBELDFTLGNLSFGWVROILQKYSVSVSAEITFTDSVYSGLPQGEAPRAQT 1094  
Qy 1080 TKVEPFEVNPPLPIVSGSSVGGILLALITAAALYKLGFFKRYQKQNMSE 1128  
Db 1095 TVLEKRYVHNPTPLIVGSSIGGLLALITAVLYKVGFFKRYQKEMEE 1143

RESULT 4  
ITAD HUMAN  
ID ITAD HUMAN  
AC Q13349; Q15575; Q15576;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).  
GN ITGAD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen.  
RX MEDLINE=96111956; PubMed=8777714;  
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,  
RA Staunton D.E., Gallatin W.M.;  
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-  
RT 3.";  
RL Immunity 3:683-690 (1995).  
RN [2]  
RP SEQUENCE OF 1-235 FROM N.A.  
RX MEDLINE=20187620; PubMed=10722744;  
RA Noti J.D., Johnson A.K., Dillon J.D.;  
RT "Structural and functional characterization of the leukocyte integrin  
RT gene CD11d. Essential role of Sp1 and Sp3.";  
RL J. Biol. Chem. 275:8959-8969 (2000).  
RN [3]  
RP SEQUENCE OF 457-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.  
RX MEDLINE=96257236; PubMed=8666289;  
RA Wong D.A., Davis B.M., LeBeau M., Springer T.A.;  
RT "Cloning and chromosomal localization of a novel gene encoding a human  
RT beta 2-integrin alpha subunit.";  
RL Gene 171:291-294 (1996).  
RN [4]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99059842; PubMed=9841932;  
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,  
RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
RT "alpha2beta2 integrin is expressed on human eosinophils and functions  
RT as an alternative ligand for vascular cell adhesion molecule 1  
RT (VCAM-1).";  
RL J. Exp. Med. 188:2187-2191 (1998).  
RN [5]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99370002; PubMed=10438935;  
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
RT binding interface between I domain and VCAM-1.";  
RL J. Immunol. 163:1984-1990 (1999).  
CC -|- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND  
CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-  
CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENESECENT ERYTHROCYTES  
CC FROM THE BLOOD.  
CC -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D  
CC ASSOCIATES WITH BETA-2.  
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -|- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL  
CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON  
CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN  
CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.  
CC -|- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -|- SIMILARITY: belongs to the integrin alpha chain family.  
CC -|- SIMILARITY: Contains 1 VWFA domain.  
CC -|- SIMILARITY: Contains 7 FG-GAP repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----

DR EMBL; U37028; AAB38547.1; -  
 DR EMBL; U40274; AAB60634.1; -  
 DR EMBL; U40275; AAB60635.1; -  
 DR EMBL; U40276; AAB60636.1; -  
 DR EMBL; U40277; AAB60637.1; -  
 DR EMBL; U40278; AAB60638.1; -  
 DR EMBL; U40279; AAB60639.1; -  
 DR EMBL; U40278; AAB60638.1; JOINED.  
 DR EMBL; AF187881; AAF62875.1; -  
 DR HSSP; F11215; IABX.  
 DR HSSP; F11215; IABX.  
 DR MIM; 602453; -  
 DR GO; GO:0008305; C.integrin complex; TAS.  
 DR GO; GO:0004895; P.cell adhesion receptor activity; TAS.  
 DR GO; GO:0015337; P.cell-cell adhesion; NAS.  
 DR GO; GO:0007160; P.cell-matrix adhesion; NAS.  
 DR GO; GO:0006955; P.immune response; NAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PRO1185; INTEGRINA.  
 DR PRINTS; PRO0453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 4.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWEA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Signal; Repeat; Calcium;  
 KW Magnesium.  
 FT SIGNAL 1 17  
 FT CHAIN 18 1162  
 FT DOMAIN 18 1100  
 FT TRANSMEM 1101 1124  
 FT DOMAIN 1125 1162  
 FT REPEAT 32 85  
 FT REPEAT ? ?  
 FT DOMAIN 150 332  
 FT REPEAT 350 400  
 FT REPEAT 401 452  
 FT REPEAT 454 516  
 FT REPEAT 518 576  
 FT REPEAT 581 633  
 FT CA\_BIND 465 473  
 FT CA\_BIND 530 538  
 FT CA\_BIND 593 601  
 FT SITE 1127 1131  
 FT DISULFID 67 74  
 FT DISULFID 106 124  
 FT DISULFID 655 710  
 FT DISULFID 769 775  
 FT DISULFID 846 861  
 FT DISULFID 994 1018  
 FT DISULFID 1023 1028  
 FT CARBOHYD 59 59  
 FT CARBOHYD 87 87  
 FT CARBOHYD 99 99  
 FT CARBOHYD 391 391  
 FT CARBOHYD 691 691  
 FT CARBOHYD 733 733  
 FT CARBOHYD 873 873  
 FT CARBOHYD 957 957  
 FT CARBOHYD 1046 1046  
 FT CONFLICT 500 500  
 FT CONFLICT 515 518  
 FT CONFLICT 825 825  
 FT CONFLICT 984 984  
 FT SEQUENCE 1162 AA; 126085 MW; P296A1A35455D77D CRC64;

Query Watch 57.8%; Score 3401; DB 1; Length 1162;

Best Local Similarity 59.3%; Pred. No. 1.5e-219;

Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

QY 1 ENLDTENAMTFOENARGFGQSVVQLOCSRVVVGAPQBIIVAAORGSLYQCDYSTGSCPEI 60  
 DB 18 ENLDEREPTTFQEDAGGFGQSVVQFGSRLVVGAPLEVAANQGRLDYDCAAAAGMCOPI 77  
 QY 61 RLOVPVEAVNMSLGLSLAAATPPQLLACGPTVHTQCTSENTYVXGLCFGLFSNLRQOPQK 120  
 DB 78 PLHIRPEAVNMSLGLTLAASNGSRLLACGPTLHRVCGENSYSKXGSCULLGSRW-EIIQT 136  
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKEWSTVWEOLKSKTLFSLMOYSEFP 180  
 DB 137 VPDATPECPHQEMDIVFLIDGSGSIDQDNFQMGFVQVAVNGQFEGDTTLFALQSVLL 196  
 QY 181 RIHTFKEFQNNPNRSLIKPIITOLLGRTHATGLRKVVRELFINITNGARKNAKIFLL 240  
 DB 197 KIHFTTFQRTSPSQOSIVDPVQLKGLTPTATGILVTVTOLPHKNGARKSANKLIVI 256  
 QY 241 TDGKFGDPLGYBDVIPDLREGVIRVIVGVGDAPFRSEKROBLMTVASKPRPRHVQIN 300  
 DB 257 TDGQKYDPLEYSDVIPQAEKAGIRYAGVGHAFQGPARTAQELMTISSAPPQDHFVKVD 316  
 QY 301 NFEALKTIQNLREKIPIAIBGTQTGSSSPHEHMSQGFSAATNSGNPLLSVGSYDWAG 360  
 DB 317 NFAALGSIQKQKELIYAVEGTQSRASSFOHENSQGFSTALTWDLGLFAGVSPSWG 376  
 QY 361 GVFLYTSKEKSTPIMNTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420  
 DB 377 GARLYPPNMSPTFINMSQENVMDRDSYLVGYSTELALWGVQLVGLGAPRYOHIGKAVIF 436  
 QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVILGAPHYTBTGGGVSVCP 480  
 DB 437 QVSRQWAKAEVTGTQISYFGASLCSVDVDSNGSTDLILIGAPHYTBTGGGVSVCP 496  
 QY 481 PRGORARWQCDVILYGEQGPWGRFGAALTALVLDGVNGDKLTDVAIGAPGEEDNRGAVYLP 540  
 DB 497 PRGORVQWQCDVILYGEQGPWGRFGAALTALVLDGVNEDKLDVVAIGAPGEENRGAVYLP 556  
 QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQBLTMDGLVDLTVGAQHVLLESQ 600  
 DB 557 HGASEGSGISPSHSQRIASSQLSPQLYFGQALSGGQBLTQDGLMDLAVARGQVLLKSL 616  
 QY 601 PVLVKALMBENPREVARNVFECDQVVKGEAGEVERVCLHVQKSTRDRLEGOQTQSVVT 660  
 DB 617 PVLKGVANRSPVEVAKAVYRCWEBEKSALAEAGDATVCLTIQKSSLDQL--GDIQSSVR 674  
 QY 661 YDLALDGRPHSRVAFVNETKSTRTQVGLQATCTETLKLQKNCIEDPVSPIVLRNF 720  
 DB 675 PDLALDGRLTSLRAIFNETKNPTLTRBKTLGLGHCHCETLKLLEDCVEDVVSPIHLNF 734  
 QY 721 SLVGTPLSAFCNLRPLAEDAQRLEPTALFPEKKNQDNICODDLSTIFSPMSLDCLVVG 780  
 DB 735 SLVREPSPQNLPRPLVAVGQDLPTASLPFERKNCQDGLCEGDLGVTLSPSGQLTVG 794  
 QY 781 GPREFNVTVYRNDGDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACRSASSTEV 840  
 DB 795 SSLEINVI TVVNAGEDSYGTWVSLYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 853  
 QY 841 SCALXSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTPEF 900  
 DB 854 SG-LRSSRCSVNHPIFHEGSGNGTIVTFDVSFKATLGRMLMRASSENKAKSSSKATP 912  
 QY 901 QLELPVKYAVVMVVTSHGVSTKYNF-TASENTSRVMQHQYQVSNLQORSPLISLVFLVP 959  
 DB 913 QLELPVKYAVVMVVTSHGVSTKYNF-TASENTSRVMQHQYQVSNLQORSPLISLVFLVP 972  
 QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELAKAPVWNCISAVCORIQCDI 1019  
 DB 973 VLLNGVAVMDVMBAPSQSL--PCVSEKPPQHSDFLTQTSRSPMLDCSIDCLQFRCDV 1030  
 QY 1020 PFFGIGQEFNATLKNLSFQWYIKTSHNHLIIVTAEILFNDTSVFTLLFGQAFVRSOTE 1079  
 DB 1031 PPSVQELDLFTLKNLSFQWVRETLOKVLVSVVAEITPDTSVYSQLPQGEAFWRAQWE 1090  
 QY 1080 TKVBPFEVNPFLPLIVGSSVGLLILLALLITLAALYKLGFEFRQYKDMNSE 1128



```

FT CARBOHYD 65 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 89 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 188 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 649 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 670 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 726 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 730 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 862 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 885 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 897 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1060 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1071 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARSPLIC 954 Q -> QGVHGLVEMQTSKILCRPAGDAEHTVGAQGEPLCP
      PMSVSEAFNDNRAGPCR (in isoform 2).
      /FTIG-VSP 002738.
      R -> W (IN REF. 1 AND 2).
      Y -> I (IN REF. 2).

FT CONFLICT 214
FT STRAND 660
FT STRAND 155
FT STRAND 164
FT TURN 166
FT HELIX 185
FT TURN 186
FT STRAND 191
FT STRAND 202
FT HELIX 208
FT HELIX 217
FT TURN 222
FT STRAND 229
FT HELIX 233
FT TURN 244
FT HELIX 247
FT TURN 250
FT TURN 253
FT STRAND 256
FT HELIX 274
FT TURN 277
FT STRAND 280
FT HELIX 288
FT HELIX 293
FT TURN 298
FT HELIX 300
FT HELIX 307
FT STRAND 311
FT TURN 318
FT HELIX 319
FT TURN 329
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match
Best Local Similarity 34.3%; Pred. No. 2.1e-95; Length 1170;
Matches 403; Conservative 210; Mismatches 460; Indels 101; Gaps 37;

QY 1 ENLDTENAMTFQ--ENARFGQSVVCLGSRVVGAPQEIIVAAORGSLYQCDYSTGSC 58
DB 26 YMLDVGRASFPFPAGRGHFGYRVLOV--GNGVIVGAPGE---GNSGSLYQCSGTGHL 81
QY 59 PIRLOQVEBAVNMSLGLSLAATSPQLLACQPTVHQTCTSENVTYKGLCFPGSNLR--- 115
DB 82 PVTLR--GSNVTSKYLQMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGP 140
QY 116 -CQPKQFPFALRGCPQEDSDIAFLDGSIIIPDPRMKWVSTWMEOLKSKTLFSLM 174
DB 141 LQRPFGQFCIKG----NVDLVFPGSMLSPDFEQKILDFMKVKKLSNITSYQFAV 196
QY 175 QYSEEFRIHFTKEFQNNPNRSLXIPITQLGRTHATGLRKVVRFLNITNGARKONAF 234
DB 197 QFSTSYKTEFDGSDYVKRDPDALLKHVKHMLLINTFGAINVATEVFEELGARPAT 256
QY 235 KILFLITDEKFGDPLGYEDVPELDRGVIRVIVGVGAPFSEKSRQRLNTVASKPPED 294
DB 257 KVLIIITDGE--ATDSGNIDAADK-----IIRYIIGIKGHFOTKESQETLHKFASKP 309
QY 295 HVFQINNFEALTKIQNLKEKIPAEIBGTQTGSSSSPEHEMSQEGFSAATISNGPLISTVG 354

```

## RESULT 6

```

ITAL_MOUSE
ID ITAL_MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain)
DE (CD11a).
GN ITGAL OR LFA-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Db 310 FVKILDTFEKLKOLFTELQKIYVIEGTSKQDLTSFNNELSSSGISADLSRGHVAWGAVG 369
QY 355 SYDNAGGVF--LYTSKEKSTPINMTRVDSMDNDAYLGYAAA--IILNRVQSLVLAGAPYQH 412
DB 370 AKDWAGFLDLKADLQDDTFIGNEPLTPEVRAGLYGTVTWLPSRKQKTSLLASGAPYQH 429
QY 413 IGLVAMFR--ONTGWESNANVKTQIGAFGLASLCSVDVDSNGSTDLVLIGAPHYEQT 470
DB 430 MGRVLLPQEPQGGHWSQVQTHGTQIGSYFGBELGVDDVDQGETELLGLLGLPFGYEQ 489
QY 471 RGGQVSVCLPRGORARWQCDAV--LYGEOQPMGRFGAALTVLGCVNGDKLTDVAIGAP 528
DB 490 RGEVPIY-----QERQLGFEVSELOQDFGYPGRFGAALTALTDINGDLVDVAGAP 544
QY 529 GEDNRGAVLYFHTSGSGISPSHSQRIAGSKSLPRLOYFGQSLSGQDLTMDGLVLTIV 588
DB 545 LEE--QGAVYIFNGRHG--GLSPQPSQRIEGTQIGSYFGBELGVDDVDQGETELLGLLGLPFGYEQ 601
QY 589 GAQGHVLLLSQPVLRVKATMEFNPREVARNVPCNDQV--KGKEAGEVRVCLLVQKSTR 647
DB 602 GABSQMIVLSRPPVDMVTLSFSFAPIPVHEVCSTYSNKMKEGVNITICFOI--KSLY 660
QY 648 DRLEEGIQSVVYDLDLDSGRPHSRAVFNKTNSTRQTVLGLTQTCETLKLQLPNCI 707
DB 661 PQF--QGRLVANLTYTLQDGHRTRRRGLFPGGRHELARNIAVT--TSMSCITDFSHPFVCV 718
QY 708 EDPVSPIVLRLNPSL---VGTPLS--AFGN-----LEPVLAEQAQLFTALFPKKNCGN 757
DB 719 QDLSPINVSINFLNWBEEGTPDRQAGQKIDIPILAPSLHSETWEI-----PEKNCGE 773
QY 758 DNICQDDLSTIFFSMSLDCLVWGGPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVS 817
DB 774 DKCEANLRVSFSFARSALRLTAPASLSVELSLNLEEDAYWYQLDLHPFPGLSFRKVE 833
QY 818 TLQNRQORSWRPLACES--ASSTEVSGALKSTSCINHPPIPPENSEVTFNITFDVDSKAS 875
DB 834 ML---KPHSQIPVSCBELPESRLLSAL---SCNVSPIPFKAGHSVALQMMFNTLVNS 887
QY 876 LGNKLLKANVTSENN---MPTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNPFTASEN 931
DB 888 WEDSVLEHNAVTCNNEEDSLLEDNSATTI---IFILYPINILIQDEBETLVVSTPKGP 944
QY 932 TSVVQHQYQV---SNLQGRSLP--ISLVFLVPRLNQTVWDRPQVTFSENLSTCHTK- 986
DB 945 KIHQVKMYQVRIQPSIHDHNIPTLEAVGVGPQPEGPITHQWSVQVMEPPV--PCHYED 1002
QY 987 -ERLPSHSD--FLAEERKAPVWNCISIAVCQRIQCDIPFFGIQEBFNATLKNLSFDWYIK 1043
DB 1003 LERLPDRAAPCLPGALFPCPVV-----PRQILVQVIGTLELVGEIE 1044
QY 1044 TSHNELIVSTABILFNDSVFTLLPGQAFVRSQTKVPEFVFNPLPLIVGSSVGGLL 1103
DB 1045 AS-SMPSLCSLSISFNSSKHFHLYGSGNASL-AQVMKVDVVYKQMLYLVLSGIGLL 1102
QY 1104 LLALITAILYKLGRKQYKQMMSEG--GPPGAEP 1136
DB 1103 LLLLIPTVLYKVGPFKNLKEKMEAGRGVPGNP 1136

```





FT	CARBOHYD	1013	1013	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1055	1055	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1086	1086	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1167	AA; 128983 MW; B0331C115DCCCFD CRC64;	
Query Match				
Best Local Similarity 28.8%; Pred.No.116-68; Indels 205; Gaps 43;				
Matches 355; Conservative 214; Mismatches 459;				
QY	1	FMLDTENA--MTPOENARGFGOSVVOLOGSVRVVGAPQBIIVANORGS-----LYQCDY	52	
DB	20	ENKVDVWAWVTALQPCGAPAVLSLLHLDPSN-----NOTCLLVARESSNNTAALYRCAL	74	
QY	53	SPGSCPIRLQVPEAVNWSLGLSLAATT--SPOLLAC-GPTVHCTCSZNTYVXGLCEL	109	
DB	75	SI--SPDEIACQ--PVEHICMPKRYQGVTLVGNHNGVLVCIOQVARKFSLNSELTGACSL	132	
QY	110	FGSNLRQQQPKPEALRG-----C-----POB	131	
DB	133	LTFNLDLQAQYFSDLEGFLDPCGAHVSDGYCHSKGSGTGEEKSARRRTVEEEDDED	192	
QY	132	DSDIAPLIDGSGIIPHDPRMKWYVTVXQOL--KKSRTLFLSMQYSEBFRHFTFKPF	189	
DB	193	GTEIAIVLDGSGIGSDPKAKNFIPTWNRNPFYKCFECNFALVQYGAIVQTEFDLOS	252	
QY	190	QNNPNPSRLIKPTQLLGRTHATGLRKVVRLEPNTNGARKNAFKILFLLTDGKFGDP	249	
DB	253	RDINASLAKVQSVQVKEVTKASAMQHVLNFIIPSGSRKKALKVWVLTGDIQDP	312	
QY	250	LYVEDVIPELDREGVIRYVIGVDAPFRSEKROELNTVASKPPRDHVFOINNPEALKTQ	309	
DB	313	LNLTWTINSPKQGVVRFALGVDRPKNNNTVRELKLIASDPKEATFKVWYNSALDGLL	372	
QY	310	NOLREKIPALGQTQSGSSFFHEMSQEGFSAITSNP--LSTVGSYDWAQGVFLY--TS	367	
DB	373	SKLQQRIVHEMG--VGDALQVQLAQCTGFSQAQLDKQVVLGTGAFNWSGALLYSTQ	429	
QY	368	KEKSTFINMT--RVDS--DMNDAYLGYAAAILNRVQSLVGLGAPRYQHIGLVAMFRONTGM	425	
DB	430	NGRGCFNLQTAKEGDSRTVQVSYLGYSLAVLHKAHGTSYVAGAPRHKLRGAVPELRKEDR--	488	
QY	426	WESNA---NVKFTQIGAYFGASICSVDVDSNGSDTLVLGAGHYVETRGQGVVCPPLR	482	
DB	489	-BEDAPVRRTEGQMSYFGSVLCPVDIMDGTDFLLVAAPFYHIRGEBGRYVYQVPE	547	
QY	483	GQEARWQCDVLYGQEQFWGRFGAALTIVLGDVNGDKLTDVAIGAP-----GEEDNRGA	536	
DB	548	-QDASFLAHTLSCHFGLTNSRFGFAMAAGVDINQKFTDVAIGALEGFGAGDGASYGS	606	
QY	537	VYLFHTSGSGISPSHSORISAGSLSPRIQYFQOSLSGGQDLTMDGLVLTGQAQCHVLL	596	
DB	607	VYTYNGHSG--GLYDPSQOTIRASSVASGLHYFGMSVSGGLDFNGDGLADITVGSRDSAVV	665	
QY	597	LRSQPVLRVKAIMEFNPREVARNVFECNDVVKGKEAGEVRVCLHVOKS---TRDLREG	653	
DB	666	LRSPVPVDLTVSMTFP-----DALPMVPIGKM--DVNLCFEYDSSVVASEPGLREM	715	
QY	654	QIOSVVTYDIALDSGRHRSRAVNETNSTRTQTVLGLTQTC-----	696	
DB	716	FLNFTVDV-----TKQRLQCEDSSGQCSLRKWNKGSFLCEHFWLI	760	
QY	697	ETLKLQLPNCIEDPVSPIVLRNLFSLVGTPLSAPGNLR---PVLAEADQRLTALF--P	750	
DB	761	STEEL-----CEDCPNNTIKYVE-----PQTSGGRRDYPNTL--DHVKPSALFQLP	809	
QY	751	PERKNGNDNICQDLSITFSMGLDLCLVWGSPRENVTVTVRNDGSDSYRTQVTFPPDL	810	
DB	810	YEKDCRKNVFCIAIEIGLTTN--ISQELVGVGVTKTEVNTANISLTSNGSDSYNTNMALNPRN	868	
QY	811	LSYKYSTLQNRQSQRSLKACASASTEVSGALKSTCSINHPPIPPENSEVTNFTFDV	870	
DB	869	LOFKKI-----QKPSVPDQCCDDPFXV---ASVLWNCKKGHPIL--KGSNVVSVTWQL	918	

RESULT 8

ITAE HUMAN

ID	ITAE HUMAN	STANDARD;	PRT; 1179 AA.
AC	P38570; Q9NZ9;		
DT	01-OCT-1994 (Rel. 30, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1		
DE	antigen) (CD103 antigen) (Integrin alpha-IBL).		
GN	ITGAE.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1] SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.		
RP	TISSUE=Leukemia, and Lymphocytes;		
RC	MEDLINE=94164962; PubMed=8119947;		
RX	Shaw S.K., Cepek K.L., Murphy E.A., Russell G.J., Brenner M.B.,		
RA	Parker C.M.;		
RA	"Molecular cloning of the human mucosal lymphocyte integrin alpha E		
RT	subunit. Unusual structure and restricted RNA distribution.";		
RL	J. Biol. Chem. 269:6016-6025 (1994).		
RN	[2] REVISIONS TO 88-114.		
RP	Parker C.M.;		
RA	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.		
RL	[3] SEQUENCE OF 53-1179 FROM N.A.		
RP	TISSUE=Fetal kidney;		
RC	MEDLINE=20138496; PubMed=10673275;		
RX	Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G.,		
RA	Shotelsarsuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,		
RA	Green E.D.;		
RT	"The genomic region encompassing the nephropathic cystinosis gene		
RT	(CTNS): complete sequencing of a 200-kb segment and discovery of a		
RT	novel gene within the common cystinosis-causing deletion.";		
RL	Genome Res. 10:165-173 (2000).		
RN	[4] MUTAGENESIS OF ASP-109 AND PHE-316.		
RP	MEDLINE=20400502; PubMed=10837471;		
RX	Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,		
RA	Brenner M.B.;		
RT	"The role of alpha and beta chains in ligand recognition by beta 7		
RT	integrins.";		
RL	J. Biol. Chem. 275:25652-25664 (2000).		
CC	FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT		
CC	MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL		
CC	CELL MONOLAYERS.		
CC	SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA		
CC	SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A		
CC	DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.		

QY	871	DSKASLGNKLLKANKVTSNNMPTNTEPQLELPVAVVAVVAVVTSCHVSTKYLNTASE	930
DB	919	EESVFPNRTADITVTSNNEKSLARETR---SLQFRHAFIAVLRS--PSVYMN--TSQ	971
QY	931	NTSRVMOHOYVSNLQGRSLPISLVFLVPLVRLNQTVIWDPRQVTFSENLST-----CHT	985
DB	972	SPSDHKPEFFNVHGENLFGAVFQICVPIKQDP-----QIVRVKNTKTQDHTECTQ	1025
QY	986	KERLPSSHSDFLAELRKAPVNVCSIAVCQRIQCDIPFGIQEEFNATLKNLSFDWYIKTS	1045
DB	1026	SQEPACGSDPVQHVKEWESVVCAL-----TSNKENTVVAEISVVG	1065
QY	1046	HNHLLIVSTA-----ELFNDVSFTLLPGCGAFVRSOTETKVBPF-----EVPNPLPIV	1095
DB	1066	HTKQLLRVDSLEPILGEISFNKSLYEGNAE-----NHRKTIIVFLKBEETSLPLII	1119
QY	1096	GSSVGGLLALLALITAAALYKLOFFKQYKOMMSR	1128
DB	1120	GSSIGGLAVLVIIAILFKCGFFKRYQQLNLE	1152

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN  
 CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL  
 CC EPITHELIAL CELLS.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VMPA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd103.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L25851; AAB59359.2; -;  
 CC EMBL; AF168787; AAF43107.1; -;  
 CC PIR; A53213; A53213.  
 CC DR HSSP; F11215; LA8X.  
 CC DR Genew; HGNC:6147; ITGAE.  
 CC MIM; 604682; -;  
 CC GO; GO:000305; C:integrin complex; TAS.  
 CC GO; GO:004895; F:cell adhesion receptor activity; TAS.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC DR InterPro; IPR002035; VMP\_A.  
 CC DR Pfam; PF01839; FG-GAP; 3.  
 CC DR Pfam; PF00357; Integrin\_A; 1.  
 CC DR Pfam; PF00092; vwa; 1.  
 CC DR PRINTS; PR01185; INTEGRINA.  
 CC DR PRINTS; PR04453; VMPADOMAIN.  
 CC DR SMART; SM00191; Intc\_alpha; 3.  
 CC DR SMART; SM00327; VWA; 1.  
 CC DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC DR PROSITE; PS00234; VMPA; 1.  
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 CC Signal; Repeat; Polymorphism; Magnesium; Calcium.  
 CC SIGNAL 1 18  
 FT CHAIN 19 1179 INTEGRIN ALPHA-E.  
 FT CHAIN 19 177 INTEGRIN ALPHA-E LIGHT CHAIN.  
 FT CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.  
 FT DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1125 1147 POTENTIAL.  
 FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 181 198 GLU-RICH (ACIDIC).  
 FT REPEAT ? ? FG-GAP 1.  
 FT REPEAT ? ? FG-GAP 2.  
 FT REPEAT ? ? X-DOMAIN (EXTRA DOMAIN).  
 FT DOMAIN 145 199 VMPA.  
 FT REPEAT 200 391 FG-GAP 3.  
 FT REPEAT 401 456 FG-GAP 4.  
 FT REPEAT 457 506 FG-GAP 5.  
 FT REPEAT 510 571 FG-GAP 6.  
 FT REPEAT 573 638 FG-GAP 7.  
 FT REPEAT 641 693 POTENTIAL.  
 FT CA BIND 522 530 POTENTIAL.  
 FT CA BIND 586 594 POTENTIAL.  
 FT CA BIND 654 662 POTENTIAL.  
 FT SITE 1150 1154 GRPKR MOTIF.  
 FT DISULFID 70 73 BY SIMILARITY.  
 FT DISULFID 126 159 BY SIMILARITY.  
 FT DISULFID 706 762 BY SIMILARITY.  
 FT DISULFID 823 829 BY SIMILARITY.  
 FT DISULFID 893 907 BY SIMILARITY.  
 FT DISULFID 1008 1033 BY SIMILARITY.  
 FT DISULFID 1041 1057 BY SIMILARITY.  
 FT CARBOHYD 49 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 321 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 444 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 444

FT	CARBOHYD	726	726	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	782	782	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	857	857	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	934	934	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	954	954	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1065	1065	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1096	1096	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	360	360	D -> E.
FT	VARIANT	1041	1041	/FTID=VAR_008884.
FT	MUTAGEN	208	208	/FTID=VAR_008885.
FT	MUTAGEN	316	316	D->A: LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	477	477	P->A: LOSS OF E-CADHERIN BINDING.
FT	CONFLICT	482	482	V -> I (IN REP. 3).
FT	CONFLICT	482	482	Q -> R (IN REP. 3).
FT	CONFLICT	950	950	R -> W (IN REP. 3).
FT	CONFLICT	1019	1019	A -> V (IN REP. 3).
SQ	SEQUENCE	1179	1179	AA; 130088 MW; E558902EDFD9D95E1 CRC64;

Query Match 19.4%; Score 1140; DB 1; Length 1179;  
 Best Local Similarity 28.9%; Pred. No. 4.1e-68;  
 Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;

QY	45	GLSYQCDYSTGS--CEPI-RLQVP-----VEAVNMSIGLSLAATTSPPQLLAGCPTVHQ	95
DB	65	GPLRCSLWQDEILCHPVERHVPDKGRHGVTVVRSHGVLCI-----QVLRVP--HS	117
QY	96	TCSENVVVKGLCFPGSNLRQPPQ-----	119
DB	118	LSSELT---GTCSLLGPDLPQAQANFPDLENLLDPDARVDTGDCYSNKEGGEDDVNTA	174
QY	120	KPPEALRGCPQED-----SDIAELIDGSGSIIPHDFRMEKEWSTVMEQL--	164
DB	175	QRRALEKEEEDKEEEDDEEAEAGTATIALIDGSGSIDPDPQRAKDFISNMNRNYE	234
QY	165	KKSKTLPSLMOYSBEFRIHFTKFEQNNPNPERSIKPITQLLGRTHATGLRKVVRELFN	224
DB	235	KCFECNALVQYGGVIQTEFPDRDSQDVMSLARVQNTQVGSVTKTASAMQHVLDSIFT	294
QY	225	ITNGARKNAFKILFLTDXGKPGDPLGYEDVIPDLREGVIRYVIGUCDAFSEKSRQEL	284
DB	295	SSHGSRKASKVMVVLTDGGIFEDPLNLITVINSFMQGVVERFAIGVGEFKSARTAREL	354
QY	285	NTVASKPPRDHVQINNFEALKTIQNLREKIPAEIGTQTGSSSFSEHESQEGFSAAIT	344
DB	355	NLIASDPDETAFKVTNTWALDGLSKLYNLI SNEGTL---VGDALHYQLAQIGFSAQIL	411
QY	345	SNGP--LLSTVGSYDMAGGVFLY--TSKEKSTPFINTRVDSMDNDA-----YLVAAAILRN	398
DB	412	DERQVLLGAVGAFDWSGGALLYDTRSRGRFLNQTAADAAADAAQAQSYLGYAVAVLHKT	471
QY	399	RVOSILVLCAPRYQHIGLVAMFR--QNTGWESENAV--KGTQIGAYFGASLCSVDVDSNGST	456
DB	472	CSLSYVAGAPQYKHG--AVFELQKEGREASFLPVLEGEQMSYFSGSELCPVIDIMDGST	529
QY	457	DLVLIGAPHYEQTRGGQSVSCPLPRGARWQCDAVLYGEGQGPWGFAGALTVLGDVN	516
DB	530	DELLVAAPFYVHGSEGVVVVRLSE--QDGSFSLARILSCHHPFTNARFGFAMAAGDLS	588
QY	517	GDKLTDVAIGAP-----GREDNR--GAVYLPHGTSGSGISPSHSORIASKLSPLRYQFGQ	570
DB	589	QDKLTDVAIGAPLEGFGADQASFGSVIYING--HWDGLSASPSQIRASTVAPGLQYFGM	647
QY	571	SLSGGQDLTMDGLVDLTGAAQGHVLLLRSQPLRVKAIEMFNPREVARNVFECNDQVVKG	630
DB	648	SNAGGFDSGDGLADITVTLGQAVFISRPVRLKVSMAFTFPSALP-----IGF	697
QY	631	KEAGEVRVCLHVQKSTRDLREGQIQSVVTDALDSCPHSRVFNFTKSTRQTQVL	690
DB	698	NGVNVRLCFEI--SSVTTASBSGLEALLNFTLDVGVKQRRRLQCSQSVRSCLGCLREWS	756
QY	691	GLTQTCETLKQLPN-----CTEDPSPVPLRLNFSLVGTPLSAPGNLRPLVAEDAQRLFT	746

Db 757 SGSQLCEDL-LMPTGELCEEDCFNSASVKYSQL-QTPGQTDHPQPILODYTEPPAI 814  
Qy 747 ALPFFPKNGNDNICDDISITFSFMSLCLVVGGRPFNVTVTVNDGDSYRTQVTF 806  
Db 815 FQLPYEKACKNGLFCVAELQLA-TTVSQELVWGLTKELTNLNTSGEDSYMTSMALN 873  
Qy 807 FPLDLSVRKYSTIONQRSQSWELACESASSTVSGALKSTCSINHPIFPENSEVTFNI 866  
Db 874 YPRLQ-----LKRQKPPSPNIQDDPPV---ASVLIMCRIGHVPL-KRSSAHVSV 923  
Qy 867 TFDVDSKASLGNKLLKANVTSSN-----MPRNKTEFO---LELPVKVAVVAVTSHGV 919  
Db 924 VQOLENAPNRTADITVTVNSNERRSLANETHLQFRHGFVAVLSKPSIMYVNTGQGL 983  
Qy 920 S--TKYLNFTASENTSVMOHQOVSNLQORSLSPLSLVPLVPLVRLAQTVINDRPOVTFSE 977  
Db 984 SHKEFLFHVHGEN---LPGAETQ-----LQICVPLKGLQVAAVKLTTRQ 1028  
Qy 978 NESTCHTKERLPSHSDPLAELRKAPVYVNCIAVCQRIQDIPFGIQEENFATLKNLS 1037  
Db 1029 ASVTCTWSQERACAYSS-VQVHEEMHSVSCVIA-----SOKENVTVAAIS 1073  
Qy 1038 PDWYKTSNHLIVST-----AEILFNDVSFTLLPGQAFVRSQTEIKVEFEVFNPL 1091  
Db 1074 WD-----HSEBLKDVTELOILGEISFNKSLYEGLNAMH--RTKITVTVFLDKYHSL 1125  
Qy 1092 PLIVGSGVGLLILALITAAALYKLGFPKQYQDMSE 1128  
Db 1126 PIIKSGVGLLVILVILVILKPGFPKRYQQLNLB 1162  
RESULT 9  
ID ITAL HUMAN STANDARD; PRT; 1151 AA.  
AC P56199;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Integrin alpha-1 (laminin and collagen receptor) (VLA-1) (CD49a).  
GN ITGAL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93155124; PubMed=8428973;  
RA Briesewitz R., Epstein M.R., Marcantonio E.E.;  
RT "Expression of native and truncated forms of the human integrin alpha  
1 subunit."  
RL J. Biol. Chem. 268:2989-2996(1993).  
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-  
B-R IN COLLAGEN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1  
ASSOCIATES WITH BETA-1.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD49a entry;  
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49a.htm".  
CC PIR; A45226; A45226.  
DR PDB; 1QC5; 17-MAY-00.  
DR Genew; HGNC:6134; ITGAL.  
DR MIM; 192968; -;  
DR GO; GO:0008305; C:integrin complex; TAS.  
DR GO; GO:0004895; P:cell adhesion receptor activity; NAS.  
DR GO; GO:0005518; P:collagen binding; TAS.  
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.  
DR InterPro; IPR000413; Integrin\_alpha.

DR InterPro; IPR002035; VWF A.  
DR Pfam; PF01839; FG-GAP; 3-  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR SMART; SM00191; inc\_alpha; 5.  
DR SMART; SM00327; vwa; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
Repeat; Calcium; Magnesium; 3D-structure.  
PT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).  
PT TRANSMEM 1114 1136 POTENTIAL.  
FT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 16 75 FG-GAP 1.  
FT REPEAT ? ? FG-GAP 2.  
FT REPEAT 147 360 VWFA.  
FT REPEAT 349 404 FG-GAP 3.  
FT REPEAT 405 457 FG-GAP 4.  
FT REPEAT 459 520 FG-GAP 5.  
FT REPEAT 540 599 FG-GAP 6.  
FT REPEAT 602 654 FG-GAP 7.  
FT CA\_BIND 470 478 POTENTIAL.  
FT CA\_BIND 552 560 POTENTIAL.  
FT CA\_BIND 614 622 POTENTIAL.  
FT SITE 1139 1142 GPFKR MOTIF.  
FT DISULFID 54 64 BY SIMILARITY.  
FT DISULFID 660 663 BY SIMILARITY.  
FT DISULFID 675 728 BY SIMILARITY.  
FT DISULFID 780 786 BY SIMILARITY.  
FT DISULFID 850 858 BY SIMILARITY.  
FT DISULFID 1002 1034 BY SIMILARITY.  
FT DISULFID 1037 1044 BY SIMILARITY.  
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 432 432 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 504 504 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 752 752 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 855 855 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 880 880 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 887 887 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 946 946 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 980 980 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;  
Query Match 18.6%; Score 1093.5; DB 1; Length 1151;  
Best Local Similarity 27.8%; Pred. No. 5.2e-65;  
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;  
Qy 1 FNLDTENAMTFQENARG-FGQSVVQL----QGRVVVVGAFQOEIVANQORSLVQCDYSTGS 55  
Db 1 FNVDVKNMTFSGVEDMFGYTVQYENEEGKWLIGSLVGGPKNKTGDVYKCPVGRGE 60  
Qy 57 CEP-IRLQVPVEA-----VNMISLGLSLAATTPPQLACGPVTHQTCSENTYVVKGL 106  
Db 61 SLEPCVKLDLPVNTSINPVTEVKENMTFGSTL-VTNPGGFLACGPLYAROGHLHYTTGI 119  
Qy 107 CFLFGSNLRQQPQKPFPEALRGCFQEDSDIAFLIDSGSIIPIHFRFRKKEWVSTVMSQLKK 166

Db 1111 RVPLWVLLSAFAGLLMLLMLLALWKGFFRPLKMK 1151

RESULT 10

ITAH HUMAN STANDARD; PRT; 1189 AA.

AC Q9UKX5; Q9UKQ1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Integrin alpha-11 precursor.

GN ITGALL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal heart, and Osteoblast;

RX MEDLINE=99411678; PubMed=1046209;

RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D., Wang S.-X., Morris C.M., Kriesansen G.W.;

RT "Cloning, sequence analysis, and chromosomal localization of the novel human integrin alpha11 subunit (ITGALL).";

RL J. Biol. Chem. 274:25735-25742(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal muscle, and Uterus;

RX MEDLINE=99395147; PubMed=10464311;

RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;

RT "cDNA Cloning and Chromosomal Localization of Human alpha(11) integrin. A collagen-binding, i domain-containing, beta(1)-associated integrin alpha-chain present in muscle tissues.";

RL J. Biol. Chem. 274:25735-25742(1999).

RN [3]

RP SEQUENCE OF 954-1188 FROM N.A.

RC TISSUE=Fibroblast;

RA Andreu N., Estivill X., Escarceller M., Sumoy L.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

CC -! FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.

CC -! SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11 ASSOCIATES WITH BETA-1.

CC -! SUBCELLULAR LOCATION: Type I membrane protein.

CC -! TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS, SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.

CC -! DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING FETAL MUSCLE CELLS (IN VITRO).

CC -! DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -! SIMILARITY: Belongs to the integrin alpha chain family.

CC -! SIMILARITY: Contains 1 VWFA domain.

CC -! SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF109681; AAF01258.1; -

CC EMBL; AF137378; AAD51919.2; -

CC EMBL; AL359064; CAB94392.1; -

CC HSSP; P17301; IAOX.

CC Genew; HGNC:6136; ITGALL.

CC MIM; 604789; -

DR GO; GO:0008305; C:integrin complex; TAS.

DR GO; GO:0004895; P:cell adhesion receptor activity; TAS.

Db 120 CSDVSPFQVNSIAP--VQECSTQ-LDIVILDGNSNIYPWDS-----VTAPLNDLLK 170

QY 167 -----SKTLPSLMQVSEFRHFTFKGQONPNPRLKIPITQLLGR-THATATGLRVV 219

Db 171 RMDIGPKQTVQVGVGVNTHFNLNKYSSTEEVLAACKIVQGGQVMTALGDTAR 230

QY 220 RELFNITNGARNAKPKILFLLDGEKFGDPLGYEDVPELDREGVIRVVGVDFAFR--- 276

Db 231 KEAFTEARGARGVKVMVIVTDGESH-DNHLKVKVQDCEDENQRFSLALGSGYNGN 289

QY 277 --SEKSRQELNTVASKPRPDHVFQINNFPALATIOQLRLEKIFAIEGTQTGSSSFHEM 334

Db 290 LSTEFVBEIKSIASEPTEKHPFNVSDELALVTIVKTLGERIFALEATADQSAASFEM 349

QY 335 SOEGFSAAITSGPILLSTVGSVDAGGVFLYTSKE-----KSTP-INNTRVDSNDAYL 388

Db 350 STGFSAHYSQDWMLGAVGDMGTVMVQKASQIILPNTPIFVSTKKNPL-ASYL 408

QY 389 GYAAAILLENRYQSL-VLGAPRYQHIGLVAMFRONTGMWESNANYKGTQIGAYFASICS 447

Db 409 GYVNSATASSGDVLYIAQPRYNHTGVIIYRMEDGNIKILQTLSGEQIGSYFGSILTT 468

QY 448 VDVDSNGSTDVLIGAPHY-----YEOTR-GGOVSVCPPLPRQARWQD 491

Db 469 TDIDKDSNTDILLVGAPMTGTEKEBQGVYVALNQTIFYOMSLPIKQTCCSRQHN 528

QY 492 AVLYGEQGPWG-RFGAALTVLGDVNGDKLTDVAIGAPCEENRGAVLPHSTSGSISP 550

Db 529 SCTTENKNEPOGARFGTAAVAKOLNDGFDIVIGAPLEDHGGAVYIHG-SGKTIRK 587

QY 551 SFSORIAGSKLPRLOYFGQSLGGODLTMDGLVLTGCAQGHVLLRSQPLRVKATME 610

Db 588 EYAOIRPSGGDKTLKFFQGSIHGMDLNGDGLTDTIGGLGAALFMSROVAVVKTWN 647

QY 611 FNPREVARNVFECNDQVVGKEAG--EVRAVLHVQ-KSTRDLRLEQIQSVVYTDLALDS 667

Db 648 FEPKNVNIQKNCH---MEGKETVCINATVCEVKLSKEDIYEADLQ----YRVTLDS 700

QY 668 GRPHSRAVNET-----KNSTRQTOVLGLTQCTETLKLQLPNCIEDPVSVIVRLAPS 721

Db 701 LRQISRSFSGTQERKVQVNIYRKSEC-----TKHSFYMLDKDHPQDSVR---ITLDPN 752

QY 722 LVGTPLSAFGLNRPVLAEDAQRLFTALPFPEKNCNDNICODLSITPFSMLDCLWVG 781

Db 753 LT-DENG-----PVLDDSLPSVHEIYFPAKCGCKEKCIDLSLVAATTEKOLLIVRS 806

QY 782 PRE-FNVTVTVRNDGSDYRQVTFPPFLDLSYRKVSTLQNSORSWKLACASSTEV 840

Db 807 QNDKPNVSLTVKNTKDSAYNTRTIVHSPNLVFSGIEALQKQ-----SCSEN----- 853

QY 841 SGALKSTSCSINHPIFPENSEVTENTEDVDKSLGN-KLLKXNVTSENNMPRTNKT 899

Db 854 ----HNITCKVGYPFTRRGEMVTKILFQNTSYLMENVTIVLSATSDSEEPETLSDNV 909

QY 900 FOLELPVKYAVVTVSHGVSTKYNLFTASENTSRVMQHYQVSN-----LGORS----- 949

Db 910 VNISIPKVEVGLQFYS-SASEVHSIAANETVPEVINSTEDIGNEINIFYLIRKSGSP 968

QY 950 ----LPLSLVF-----LUPVRLNQTIVDRPQVTSNLSSTCHTKE----- 987

Db 969 MPBLKLSISFPNMNTSGNYPVLYPTGLSS-----SENACPHIFDPFSPINSOK 1017

QY 988 RLPSHSDFLAEILKAPVWACSAVACQRIQCIDIPFFGIOE-----BFN 1029

Db 1018 KMTTSTD---HLKRGVILDCNTCKPATITCNLTSSDISQVNVSLILKRPFIKSYFSSLN 1074

QY 1030 ATLKGVLSDWYTKTSHNHLILVSTAEILLFNDVSFTLLPQCGAFVRSQTETKVEPPEVN 1089

Db 1075 LATIRGEL-----RSENASLVSSSN-----QKRELAIQISDKGLPG 1110

QY 1090 PLPL--LVGSGVGLLALLITALYKLGFFKQYKDMASE 1128



Db 1048 PVE--EDLRAPQLNHSNSDVVSINCNIRLVP-NQENFHLGLN---WLSRLKLYKS 1101  
 Qy 1047 NELLIVSTAEILFNDSVFTLLPQCGAFVRSQTETVSEFEVFN-----PLPLVGGSSVG 1100  
 Db 1102 MKIMVNAALQROFH-SPF-----IFREDEPSQIVFBISKQEDWQVPIWIIVGSTLG 1152  
 Qy 1101 GLLALALTAALYKLGFTK--ROYKDMMSBEGPPQABP 1136  
 Db 1153 GLLALALVLALWKLGFERSARRRE-----PGLDLP 1183

## RESULT 11

IT22 BOVIN STANDARD; PRT; 1170 AA.  
 AC P53710;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)  
 GN (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).  
 OS Bos taurus (Bovine).  
 OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94193647; PubMed=75111592;  
 RA Kamata T., Puzon W., Takada Y.;  
 RT "Identification of putative ligand binding sites within I domain of  
 integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";  
 RL J. Biol. Chem. 269:9659-9663 (1994).  
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT  
 RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN  
 COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE  
 EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 EXTRACELLULAR MATRIX.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
 ASSOCIATES WITH BETA-1.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L25886; AAB59255.1; -  
 CC PIR; I45914; I45914.  
 CC KSSP; P17301; LAOX.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP; 3.  
 CC Pfam; PF00357; Integrin\_A; 1.  
 CC Pfam; PF00092; vwa; 1.  
 CC SMART; SM00191; Int\_alpha; 5.  
 CC SMART; SM00327; VWF; 1  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC PROSITE; PS00234; VWFA; 1.  
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 CC Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.  
 CC NON\_TER 1 1

FT SIGNAL <1 18  
 FT CHAIN 19 1170  
 FT DOMAIN 19 1121  
 FT TRANSMEM 1122 1143  
 FT DOMAIN 1144 1170  
 FT REPEAT 34 92  
 FT REPEAT ? 367  
 FT DOMAIN 177 367  
 FT REPEAT ? 475  
 FT REPEAT 423 538  
 FT REPEAT 540 599  
 FT REPEAT 604 656  
 FT CA\_BIND 488 496  
 FT CA\_BIND 552 560  
 FT CA\_BIND 616 624  
 FT SITE 472 474  
 FT SITE 1146 1150  
 FT DISULFID 72 81  
 FT DISULFID 669 726  
 FT DISULFID 778 784  
 FT DISULFID 854 865  
 FT DISULFID 1008 1039  
 FT DISULFID 1044 1049  
 FT CARBOHYD 94 94  
 FT CARBOHYD 101 101  
 FT CARBOHYD 332 332  
 FT CARBOHYD 421 421  
 FT CARBOHYD 449 449  
 FT CARBOHYD 464 464  
 FT CARBOHYD 688 688  
 FT CARBOHYD 748 748  
 FT CARBOHYD 945 945  
 FT CARBOHYD 1063 1063  
 FT CARBOHYD 1070 1070  
 FT VARIANT 580 580  
 FT VARIANT 588 588  
 FT VARIANT 725 725  
 SQ SEQUENCE 1170 AA; 128929 MW; EECFIC5F2448FB1 CRC64;

Query Match 18.2%; Score 1071; DB 1; Length 1170;

Best Local Similarity 27.4%; Pred No. 1.7e-63;

Matches 333; Conservative 219; Mismatches 435; Indels 168; Gaps 47;

Qy 1 FNLDENAMTFQ-ENARGFGOSVQL---QGSRRVVVGAQPEIVAANGRSGSYQC--DYST 54

Db 19 YNVLPRKAKIFSGPSSEQFGYAVQOFINPKGNWLLVGSFPMSPGPKNRMGDVYKCPVDLST 78

Qy 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATTSPOLLACGPTVHCTCSENVYKGLC 107

Db 79 TTCEKLNLTSTMSNVNTEKTNMSLGLTTRNVGTGGFLTCGLPMAQQCCSQYTTTGV 138

Qy 108 PLFGSNLRQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKEMVSTVMEQLK-- 165

Db 139 SDVSPDF-QLRTSPAPAVQTCF-SFIDVVVVCDESNSIYPWD--AVKNLEKFFVQGLDIG 194

Qy 166 KSKTLFSLMQVSEFRIHFTFKFQNNPNRSLIKPTQLL-----GRTHATGLKRVRE 221

Db 195 PTKQMGLIQVANNRPNVFNLTFSKD---EMIKATSTQFYQGDLTNTFKAIQADST 251

Qy 222 LFNITNGARKNAFKILPLLTDGKFGDPLGVEDYVPELDREGVIRYVIGV-----GDAFR 276

Db 252 AYSTAAGRPQATKVMVVTDGESH-DGSKLKAVIDQCNKNILRFGIAVLGYLNRALD 310

Qy 277 SEKSRQELNTVASKPPRDHVPQINNFEALKTIONOLREKIFAIKRGTTQSGSSSEHMSQ 336

Db 311 TKNLIKETKAIASIPTEHPFNVSDEADLLEKAGTIGEQIFSIKGTVOG-GDNFQMSQ 369

Qy 337 EGFSNAIT--SNGPLISTVGSYDWDAGVFLVTSKESKFINMT--RVDSDMN-DAYLCYA 391

Db 370 VGFSAEYSPQNNILMLGAVGAYDMSGTVQKTPHGLIFSKQAEQIILQDRNHSYLGYS 429

Qy 392 AAIIILNRFVQSLVLGAPRYQHIGLVAMPFRONTGMWESNANV-----KGTQIGAYFGASL 445

Db	430	VASISTGNSVHFVAGAPRANNTQIVLYSVN-ENGNTVVIQSGDQIGSYFGSVL	484	RX	MEDLINE=94363406; PubMed=8081889;
Qy	446	CSVDVDSNGSTDLVLCAPHYEQTE--GQSVSCVPLPRQARWCCDAVLXGEQOPWG	503	RA	Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,
Db	485	CAVDVNDITDVLVAGPYNNDLKEBGRVFLVITKG--ILNWH--QFLGPGNLENA	541	RA	Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.,
Qy	504	RFGAATVLDVNGDKLTDVAIGAPGEENRGAVYLFHGTSGSGISPSHSQIAGS--KL	561	RT	"The mouse VLA-2 homologue supports collagen and laminin adhesion but
Db	542	RFGSAAALSDINDGDFNDVIGSPLENQNSGAVIYNGHEG--IRLVSKILGSDRAF	600	RL	not virus binding."
Qy	562	SPRLQYFGOSLSGQDLTMDGLVLTGAGHVLRLSOPVLRLVKALMEFENPREVARNVF	621	RN	Cell Adhes. Commun. 2:131-143(1994).
Db	601	SSHLQYFGRLEDYGDNGSDITFVSGAPGVQVQVQWQSADIADVSDASTPKKI--TL	658	RP	SEQUENCE OF 450-1178 FROM N.A.
Qy	622	ECNDQVVKGEAGEVRVCLHVQKSTDRRLRBEQIQSVVTVYDLDL-----SGRPHSRAVFN	677	RC	TISSUE=Lung;
Db	659	NKXAEI-----KLXCLP-----SARFPTNQNNQVAIVNITIDEDQSSRVISRGLEK	707	RE	MEDLINE=94355691; PubMed=7521231;
Qy	678	ETKNSVTRQTVLGLTQCE--TLKQLPNCIEDPSPVILRLNPSL--VGTPLSAFGLN	733	RA	Wu J.B., Santoro S.A.;
Db	708	ENNERCLOKTMIVSQORCEYIIHQEPB---DIISPLNLCNWSLENPGT-----	756	RT	"Complex patterns of expression suggest extensive roles for the alpha
Qy	734	RPVLADQAQLFTALPFERKNCNDNICODLSITF-----SFMSLDCLVVGPGPRENFVV	789	RL	2 beta 1 integrin in murine development."
Db	757	NPALBAYSETVKVPSIPPHKDCGDDGVCISDLVNVQQLPATQQPFIYSNQNKRLTFSV	816	CC	Dev. Dyn. 199:292-314(1994).
Qy	790	TVRNDGEDSVRTQVTFPPFLDLISYRKVSTLONORSORSHLACESASST-EVSGALKSTFS	848	CC	RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
Db	817	QLKXKKEASVNTIIVDFSENLPF-----ASWMPVDGTEVTCQIASSQKSVT	864	CC	COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
Qy	849	CSINHPHPPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRTNKTFOLELPVKY	908	CC	FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
Db	865	CNVGYPAKSKQVTFINEFDNLQ--NLQNAQISPRALSESQENWADNSVNLKLSLIV	923	CC	EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
Qy	909	AVYMWVTSQVSKYLNINFTASENTRVMQHYQVSNLQOR-----SLPISLVFLV	958	CC	C-PROTEPTIDES AND E-CADHERIN. MICE HOMOTYZOUGS FOR A NULL MUTATION
Db	924	DABIHIT--RSTNINFEVSLDGNVSVV--HSPE--DIGPKFIFSIKVTTCGVSFVNA---	976	CC	IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
Qy	959	PVRLNQTIVNDROPVTFSEN-----LSSTCHTKB-----RLPHSDFLAB-	998	CC	- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
Db	977	-----SVIIHIFQYTKDKRPLMVLGVHTDQAGDISCEAEINPLKIGQTSSSVSFKSEN	1030	CC	ASSOCIATES WITH BETA-1.
Qy	999	LRKAPVNVCSIAVCQRIQCDIPFGIQQEFNATLKNLSFDWIKTSHNHLIVSTARI-	1057	CC	- SUBCELLULAR LOCATION: Type I membrane protein.
Db	1031	FRHKELNCRCTAGSCNIMCMLRDLQVKGFFLVNSTRIMNGTFAASTFTQVQLTAAAEID	1090	CC	- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
Qy	1058	LFNDSVFTL-----LPQGAFAVRSQTETKVEPFE--VPNPPLIVGVSSVGGHLLALITA	1110	CC	WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
Db	1091	TYNPQIYVISENTVTP-----LTIMPKPHEKVEVPTGVIVGVSIAGILLLLALVA	1140	CC	- SIMILARITY: Belongs to the integrin alpha chain family.
Qy	1111	ALYKLGFFKQYKDM 1125		CC	- SIMILARITY: Contains 1 VWFA domain.
Db	1141	ILWKLGFFKQYKDM 1155		CC	- SIMILARITY: Contains 7 FG-GAP repeats.
RESULT 12					-----
ID	ITA2 MOUSE	STANDARD;	PRT;	1178 AA.	EMBL; Z29987; CAA82877.1; "
AC	Q62469; Q62163;				EMBL; X75427; CAA53178.1; "
DT	15-JUL-1998 (Rel. 36, Created)				PIR; S44142; S44142.
DT	15-JUL-1998 (Rel. 36, Last sequence update)				HSP; P17301; LAOX.
DT	28-FEB-2003 (Rel. 41, Last annotation update)				MCD; MGI:96600; Itga2.
DE	Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)				InterPro; IPR000413; Integrin_alpha.
DE	(Collagen receptor) (VLA-2 alpha chain) (CD49b).				InterPro; IPR002035; VWP A.
GN	ITGA2.				Pfam; PF01839; FG-GAP; 3.
OS	Mus musculus (Mouse).				Pfam; PF00357; Integrin_A; 1.
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				Pfam; PF00092; vwa; 1.
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				SMART; SM00191; Int_alpha; 4.
OX	NCBI_TaxID=10090;				SMART; SM00327; VWA_1.
RN	[1]				PROSITE; PS00242; INTEGRIN_ALPHA; 1.
RP	SEQUENCE FROM N.A.				PROSITE; PS50234; VWFA; 1.
RC	STRAIN=C57BL/6 X CBA; TISSUE=Lung;				Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;

RA	Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;				
RT	"The mouse VLA-2 homologue supports collagen and laminin adhesion but				
RL	not virus binding."				
RN	Cell Adhes. Commun. 2:131-143(1994).				
RP	SEQUENCE OF 450-1178 FROM N.A.				
RC	TISSUE=Lung;				
RE	MEDLINE=94355691; PubMed=7521231;				
RA	Wu J.B., Santoro S.A.;				
RT	"Complex patterns of expression suggest extensive roles for the alpha				
RL	2 beta 1 integrin in murine development."				
CC	Dev. Dyn. 199:292-314(1994).				
CC	RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO				
CC	COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,				
CC	FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED				
CC	EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN				
CC	C-PROTEPTIDES AND E-CADHERIN. MICE HOMOTYZOUGS FOR A NULL MUTATION				
CC	IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.				
CC	- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2				
CC	ASSOCIATES WITH BETA-1.				
CC	- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS				
CC	WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.				
CC	- SIMILARITY: Belongs to the integrin alpha chain family.				
CC	- SIMILARITY: Contains 1 VWFA domain.				
CC	- SIMILARITY: Contains 7 FG-GAP repeats.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; Z29987; CAA82877.1; "				
DR	EMBL; X75427; CAA53178.1; "				
DR	PIR; S44142; S44142.				
DR	HSP; P17301; LAOX.				
DR	MCD; MGI:96600; Itga2.				
DR	InterPro; IPR000413; Integrin_alpha.				
DR	InterPro; IPR002035; VWP A.				
DR	Pfam; PF01839; FG-GAP; 3.				
DR	Pfam; PF00357; Integrin_A; 1.				
DR	Pfam; PF00092; vwa; 1.				
DR	SMART; SM00191; Int_alpha; 4.				
DR	SMART; SM00327; VWA_1.				
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.				
DR	PROSITE; PS50234; VWFA; 1.				
KW	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;				
KW	Platelet; Signal; Repeat; Calcium; Magnesium.				
FT	SIGNAL	1	26		BY SIMILARITY.
FT	CHAIN	27	1178		INTEGRIN ALPHA-2.
FT	DOMAIN	27	1129		EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1130	1151		POTENTIAL.
FT	DOMAIN	1152	1178		CYTOPLASMIC (POTENTIAL).
FT	REPEAT	42	100		FG-GAP 1.
FT	REPEAT	?	?		FG-GAP 2.
FT	DOMAIN	185	375		VWFA.
FT	REPEAT	?	?		FG-GAP 3.
FT	REPEAT	431	483		FG-GAP 4.
FT	REPEAT	485	546		FG-GAP 5.
FT	REPEAT	548	607		FG-GAP 6.
FT	REPEAT	612	664		FG-GAP 7.
FT	CA_BIND	496	504		POTENTIAL.
FT	CA_BIND	560	568		POTENTIAL.
FT	CA_BIND	624	632		POTENTIAL.
FT	SITE	480	482		CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1154	1158		GPFR MOTIF.
FT	DISULFID	80	89		BY SIMILARITY.

FT DISULFID 677 734 BY SIMILARITY.  
 FT DISULFID 786 792 BY SIMILARITY.  
 FT DISULFID 862 873 BY SIMILARITY.  
 FT DISULFID 1016 1047 BY SIMILARITY.  
 FT DISULFID 1052 1057 BY SIMILARITY.  
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 696 696 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;  
 Query Match 18.0%; Score 1057; DB 1; Length 1178;  
 Best Local Similarity 27.8%; Pred. No. 1.5e-62;  
 Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;  
 QY 1 FNLDTENAMTFO-ENARFGQSVQL---QGSVVVVGAPQEIIVAAANORGLYQC---DYST 54  
 DB 27 YNWGLPGAKIFGSPSEQGYVQVLTNPQGNWLLVGSFPMSPENRMGSDVYKCPVDLPT 86  
 QY 55 GSCEPRILQ-----VPVEAVNMSLGLSLAATTSPQQLACGPTWHQTCSENTYVKGILC 107  
 DB 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLTCGLPLWAHQCGNQYVATGIC 146  
 QY 108 FLFGNLRQPOQ---KPPBALRGCTQEDSDIAFLIDGSGIIPHFRMKEVWVWMEQL 164  
 DB 147 ----SDVSPDFQFLTSFSPAVQACPSL--VDVVVWCDSNSIYP--WEAVKNEFLKPVFTGL 199  
 QY 165 X--KSKTLPSLMQYSEEFRIHFTFKFQNNPNSRLIKPITQLLG-RHTATGLKRVRE 221  
 DB 200 DIGPKTKTQVALQYANEPRIIIFNLNDFETKEDVQATSETROHGGDLTWTFPAIFARDY 259  
 QY 222 LFNITNGARKNAFKILFLTLDGKFGDPLGYEDVTELPEDREGVIRYICV-----GDAPR 276  
 DB 260 AYSQTSGRFGATKVMVWVTDGESH-DGSKLKTIVIQCCNDDEILRGIAVLGYLNRNALD 318  
 QY 277 SEKSRELNTVASKPRDRHVQINNEALTKIQNLREKI FALGEGTQCTGSSSFHEMSQ 336  
 DB 319 TKNLKEIKAIATPTERTERFFNVADAEALLEGAGTLGEGQIFLEGIVQ--GNDFONERQA 377  
 QY 337 EGFS--ALTSNGPLSLTVSGVDWAGVFLYTSKEKSTFINMT--RVSDVM--DAYLGYA 391  
 DB 378 VGFSADYAFQNDILMLGANGAFDWSGLTVQETSHKPVFPKQAFDQVLQDRNHSPLGYS 437  
 QY 392 AAILLRNVQSLVGLAPRYOHIGLVAMFRONTGMWESNAV-----KGTQIGAYFGASLCS 447  
 DB 438 VAAISTEDGVHFFVAGAPRANYTQGVLYSVNK---QGNVTVIQSHRGDQIGSYFGSVLCS 494  
 QY 448 VDVSNGSLDLVLIGAPHYVEQTR--GGQVSVCPPLRGQARWQCDAVLYGQGPWGRF 505  
 DB 495 VDVKDXTITDVLVGAPTYNDILKKEGKLYLFTITKGLNCHQ---FLEGEGTGNARF 551  
 QY 506 GAALTVLGVNGDKLFDVAIGAPGEDNRGANVLFHGTSGSGISPSHRSORLAGSKLSPR- 564  
 DB 552 GSAIAALSINMDGFNDVIGSVENSGAVIYNGHQST-IRTKYSOKILGSGNGAFR 610  
 QY 565 -LQYFGSLSGGODLWDGLDITVCAQGHVLLRSCQPLVRKVAIMEFNPREVARNVPC 623  
 DB 611 HQFFGSLDGYGLNGDSITDVSIGALGVQIQLWSQSDADVVAIEALFTP----- 660  
 QY 624 NDQVVKGEAREVRLVHVKSTRDLREGIQSVVTVYDLAD----SGRPHSRVAVNET 679  
 DB 661 -DKITLNLKDAKITLCLFAEPFRPAGQNNQV--AILFNMTLDAGHSRVRTSRGVFREN 717  
 QY 680 KNSTRQTVLGTQTCET--LKLQPLNCIEDPVSVTLRLAFESLVTPLSAFGLNLRPVL 737  
 DB 718 SERFLQNVWVNVQKCEHHISIQPS---DVVNPLDLRVDLISLENPGTS-----PAL 768  
 QY 738 ABDAQRLTALPPFEKNCNGNDNICQDLSI-----TFSEMSLDCDLVWGSPREFNVT 788

Db 769 EAYSETVKVESIPFYKEGSDGICISDLILDVQOLPAIQOSP-----IVSNQKRLTFS 823  
 QY 789 VYVRNDEGDSYRTQVTFFFEFLDLSYRKVSTLQNRQSRQSWRLACESASST-EVSGALKST 847  
 Db 824 VILKNRGESAYNTVLAEPSENLF-----ASFSMPVDGTVEVCVGSQSKSV 871  
 QY 848 SCSINHIFPENSEVTNITFDVDSKASLGKLLKXANVTSENNMPRTNKT--FQLELP 905  
 Db 872 TCDVGYFALKESEQOQVTFITNFDFNLQ-NLQNAINFQAFSESO--ETNKADNSVSLTIP 928  
 QY 906 VKYAVYVAVTSHGVSTKYLNFASENTSRVMQHOYQVSNLQOR-----SLPISLV 955  
 Db 929 LLYDAELHLT-RSNTNFYEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVMA 984  
 QY 956 FLV-----PVRLNQTVINDRRQVTF--SENLS 980  
 Db 985 LVTHIPQYTKERNPLLYLTGIQTDQAGDISCTABINPLKLPHTA----PSVSPKNENFR 1040  
 QY 981 STCHTKERLPSHSDFLAELRKAPVVGSIACVQRIQCDIPFGIOEEFNATLKGNLSPDW 1040  
 Db 1041 ---HKZ-----LDCRTTSCSNITCNLKDLMKAEYFINVTVTRVWNR 1080  
 QY 1041 YIKTSHNLLIVSTABILEFNDVSFTLLPGQGAQFVRSOTETKVBSEFFVNPPLPLVGSVG 1100  
 Db 1081 FAASTFTQVLTAAAEIDTNPQLFVIEENAVTIPLMIMKTEKAEVPT--GVIIGSIIA 1138  
 QY 1101 GLALLALITRALYKLGPFKROYKDM 1125  
 Db 1139 GILLALLATAGLWKLGPFKROYKDM 1163  
 RESULT 13  
 ITAL RAT STANDARD; PRT; 1180 AA.  
 AC P18614;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)  
 DE (CD49a).  
 DE ITGAL.  
 CS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 ON NCBI\_TaxID=10116;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90338125; PubMed=2380249;  
 RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,  
 RA Esch F., Carbonetto S., Reichardt L.F.,  
 RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor  
 RT for laminin and collagen."  
 RL J. Cell Biol. 111:709-720(1990).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.  
 RX MEDLINE=99313197; PubMed=10386626;  
 RA Nolte W., Pepinsky R.B., Veenayamov S.Y., Koteliansky V.,  
 RA Gotwals P.J., Karpusas M.;  
 RT "Crystal structure of the alphabeta1 integrin I-domain: insights into  
 RT integrin I-domain function."  
 RL FEBS Lett. 452:379-385(1999).  
 CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
 CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-  
 CC E-R IN COLLAGEN.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1  
 CC ASSOCIATES WITH BETA-1.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VFMA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.



```

Db      986 YTIKRGHPMPBLQLSISFPNLTADGYVLYPIG-----WSSDNVNCRRSLDDPPG 1039
QY      972 -----QVTFP-----EMLSSSTCHTKERLPKSHSDFELAKRPVNVVNCSTIACQRIQ 1016
Db      1040 INSGKQMTISKSEVLKRGITQCSSTC-----GVAITICSLPSDLSQ 1082
QY      1017 CDI-----PFGIQEFP---NATLKGMLSPWYIKTSNHLIIVSTABILFNDVSFTLL 1067
Db      1083 VNVSLLLWKEPT-IRAFPSLMLTLRGELK-----SENSLTSSN----- 1123
QY      1068 PCGAFVVRQVETKVEPPFVNPPLP--IVGSSVGLLMLLITLALYKLGKFPKQYKQM 1125
Db      1124 -----RKRLAIQISKDGLPGRVPLWVILLSAFAGLLMLLILALWKIGFKRPLKVK 1177
QY      1126 MSE 1128
Db      1178 MEK 1180

RESULT 14
ITF2_HUMAN
ID ITF2_HUMAN STANDARD; PRT; 1181 AA.
AC P17301;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.
RC TISSUE=Endothelial cells;
RX MEDLINE=89108879; PubMed=2545729;
RA Takada Y., Hemler M.B.;
RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit
RT (platelet GP1a): homology to other integrins and the presence of a
RT possible collagen-binding domain.";
RL J. Cell Biol. 109:397-407 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Ruldane S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.
RX MEDLINE=98019223; PubMed=9353312;
RA Emsley J., King S.L., Bergelson J.M., Liddington R.C.;
RT "Crystal structure of the I domain from integrin alpha2beta1.";
RL J. Biol. Chem. 272:28512-28517 (1997).
RN [4]
RP VARIANT HPA-5 (BR).
RX MEDLINE=94043762; PubMed=7901236;
RA Santos S., Kaib R., Walke M., Kiesel V., Mueller-Eckhardt C.,
RA Newman P.J.;
RT "The human platelet alloantigens Br(a) and Brb are associated with a
RT single amino acid polymorphism on glycoprotein Ia (integrin subunit
RT alpha 2).";
RL J. Clin. Invest. 92:2427-2432 (1993).
RN [5]
RP VARIANT GLU-534.
RX MEDLINE=20206009; PubMed=10744142;
RA Kroll H., Gardemann A., Fechter A.;
RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G
RT gene polymorphism on coronary artery disease and acute myocardial
RT infarction.";
RL Thromb. Haemost. 83:392-396 (2000).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT

```

```

CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2
CC associates with beta-1. Interacts with HPSS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- POLYMORPHISM: Position 534 is associated with platelet-specific
CC alloantigen HPA-5 (BR). HPA-5A/BR(A) has Lys-534 and HPA-5B/BR(B)
CC has Glu-534. HPA-5B is involved in neonatal alloimmune
CC thrombocytopenia (NAIT or NATP). The K534B polymorphism may play a
CC role in coronary artery disease (CAD).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X17033; CAA34894.1; -.
CC EMBL; AF512556; AAM34795.1; -.
CC F01839; A33998; A33998.
CC PDB; 1AOX; 25-NOV-98.
CC PDB; 1DZ1; 02-AUG-01.
CC Genew; HGNC:6137; ITGA2.
CC MIM; 192974; -.
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0005886; C:plasma membrane; TAS.
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
CC GO; GO:0005518; F:collagen binding; TAS.
CC GO; GO:0007596; P:blood coagulation; TAS.
CC GO; GO:0007160; P:cell-matrix adhesion; TAS.
CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWFA.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF00092; Vwa; 1.
CC SMART; SM00151; Int_alpha; 5.
CC SMART; SM00327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWFA; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;
CC 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 1181
FT DOMAIN 30 1132
FT TRANSMEM 1133 1154
FT DOMAIN 1155 1181
FT DOMAIN 1155 1161
FT REPEAT 45 103
FT REPEAT 2 2
FT DOMAIN 188 378
FT REPEAT 378 433
FT REPEAT 434 486
FT REPEAT 488 549
FT REPEAT 551 610
FT REPEAT 615 667
FT CA_BIND 499 507
FT CA_BIND 563 571
FT CA_BIND 627 635
FT SITE 1157 1161
FT SITE 1157 1161

```

FT	DISULFID	83	92	BY SIMILARITY.	266	AASGRRSATKVVVVVITDGEH-DGSMKAVIQDCHNDNLRPGIAGVLGYLNRNALDTKN	324
FT	DISULFID	680	737	BY SIMILARITY.	280	SQELNTVASKPRDHPVQINNFRALKTIQNLREKIFAIEGTQGTSSSPHEMSQEGF	339
FT	DISULFID	789	795	BY SIMILARITY.	325	LIKEIKALASIPTEPFNVSDRAALLEKAGTLEQIFSIETGTVQ-GDNFQEMMSQVGF	383
FT	DISULFID	1019	1050	BY SIMILARITY.	340	SAATISNGP--LLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGAAAI	394
FT	CARBOHYD	105	105	N-LINKED (GLCNAC. . .)	384	SADYSSQNDIIMLGAAGAFGMSGTIVQKTSHGHLIPFQKQAFDQILODRHNSYLGYSVAA	443
FT	CARBOHYD	112	112	N-LINKED (GLCNAC. . .)	395	ILNRVQSLVLGAPRYOHIGLVAMFRONTGMBESNANV-----KGTQIGAYFGASLCSV	448
FT	CARBOHYD	343	343	N-LINKED (GLCNAC. . .)	444	ISTGESTHFVAGAPRANTYQGVIVSVN-----ENGNITVQARHGDQIGSYFGSVLCSV	498
FT	CARBOHYD	432	432	N-LINKED (GLCNAC. . .)	449	DVDSNGSTDLVLGAPHYEQTR--GGQVSVCPILPRQARQWQCDVAVLYGBOGQWGRFG	506
FT	CARBOHYD	460	460	N-LINKED (GLCNAC. . .)	499	DVDKDTITDVLGAPWMSDLKKEBGRVYLFITKIGILGQH---FLEGPEGIENTRFG	555
FT	CARBOHYD	699	699	N-LINKED (GLCNAC. . .)	507	AALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSORISAGS--KLSPR	564
FT	CARBOHYD	1057	1057	N-LINKED (GLCNAC. . .)	556	SAIAALSINMDGFRNDVIVGSPLENQSGAVIYNGHGT--IRTKYSOKILGSDGAFRSH	614
FT	CARBOHYD	1074	1074	N-LINKED (GLCNAC. . .)	565	LQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLLRSQPVLRVKAIMEFNPREVARNVFCN	624
FT	CARBOHYD	1081	1081	N-LINKED (GLCNAC. . .)	615	LQYFGRLDGYGLNGDSITDVSIGAPGVQVQLWSQSIADVAIEASFTEKI--TLVKN	672
FT	VARIANT	534	534	K -> E (IN ALLOANTIGEN HPA-59; dlsnp:1801106) /FTIG=VAR_003977.	625	DQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVYDLDL-----SGRPHSRVAFNETK	680
FT	TURN	170	171		673	AQ11-----LKLCF-----SAKFRPTKQNNQVAIVYNTITLDADGFSRVSRLFKENN	721
FT	STRAND	173	180		681	NSTRTOVLGLTQTC--ETLKLQPLNCIEDPVPVIVLRNLFSLVGTPLSAGNLRPVLA	738
FT	TURN	183	184		722	ERCLQKNVNVQAQSCPEHIIYIOEPS---DVNSLDLRVDISLENPGTS-----PALE	772
FT	HELIX	188	199		739	EDAQELFTALPPEKNCNDNICDDLSITP-----SFMSLDCLVVGGRPRENVTVVRND	794
FT	TURN	200	201		773	AYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSTLKNK	832
FT	STRAND	204	204		795	GEDSVRTQVTFPPFDLISYRKVSTLQNRQSRKWLACESASST--EVSGLKSTSCSINH	853
FT	TURN	206	207		833	RESAYNTGIVVDFSENLEF-----ASFSLPVDGTEVTQVAAQSKSVACDVGY	880
FT	STRAND	209	216		854	PIFPENSEVTFTNITPDVDSKASLGKLLKANVTSENNMPRTNKTETFOLELPVKYAVMV	913
FT	STRAND	220	224		881	PALKREQQVTFITNDFNLQ-NLQNASLSQALSESQENKADNLVNLKIPLLYDAEI-	938
FT	TURN	226	228		914	VTSHGVSTKYLNTASNTSRWQHQQVQVSNLGR-----SLPISLVFLV-----	958
FT	HELIX	232	240		939	---HLTRSTNINFYEISDGNVPSIVHSFEDVGPKFIFSLKVTGTSVPVSMATVLIHIPQ	995
FT	TURN	241	241		959	-----PVRLNQTVIWDPRQVTP--SENLSSTCHTKER	988
FT	HELIX	252	262		996	YTKRKNPLMYLTGVQTDKAGDISCNADINPLKIGT---SSSVSFKSENF--HTKE-	1047
FT	TURN	263	264		989	LPSSHSDFLAELRKAPVNVCSIAVCQRICQDTPPGIQEEFNATLKNLSFDWYIKTSENH	1048
FT	TURN	266	268		1048	-----LNCRATASCNVTCWLKDVHMKGEYFNVTTRINWGTFASSTFQT	1091
FT	TURN	269	269		1049	LLIVSTAEI-LPNDVSFTLLPQCGAFVRSQETKVPPEVNP-----LP--LIVGSSVG	1100
FT	TURN	275	282		1092	VQLTAAAEINTYNEIYVI-----EDNTVTIPLIMKPKDEKAEVFTGLIGSIIA	1141
FT	STRAND	289	291		1101	GLLLALLIATAALYKLGPFKRYKDM	1125
FT	HELIX	292	301		1142	GILLALLAVAILWKLGGPFKRYKDM	1166
FT	TURN	302	303				
FT	STRAND	304	311				
FT	HELIX	313	317				
FT	TURN	318	319				
FT	HELIX	323	330				
FT	TURN	331	332				
FT	HELIX	337	340				
FT	STRAND	341	344				
FT	HELIX	347	353				
FT	HELIX	354	362				
FT	TURN	363	363				
SQ	SEQUENCE	1181	AA; 129295 MW; 7E1B7ED968A94070 CRC64;				

Query Match

Best Local Similarity 17.9%; Score 1054; DB 1; Length 1181;  
Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;

QY	1	FNLDTENAMTPO-ENARFGGVSQVL---QGSRVVVGAPQEIIVANORGSLYOC--DYST 54
DB	30	YVVGLEPAKIFSGPSEQFGYAVQOQFINPKGNWLLVGSFMSGPPENRMAGDVVKCPVDLST 89
QY	55	GSCPEIRLQ-----VPVEAVNWSLGLSLAATTSPPQLACAGPTVHQTCSNTYVKGLC 107
DB	90	ATCEKLNLTSTSIINVTETKNNLSGLILTRNNGTGGFLTCGLPWAQQCGNYTTGVC 149
QY	108	PLFGSNLRQQPKFPEALRGCPQEDSDIAPLIDGSGIIPHDFRMKEWSTVMEQLK-- 165
DB	150	SDISDPD-QLSASFSPATQCPESL-IDVVVVCDESNSIYPWD--AVKNFLEKTVQGLDIG 205
QY	166	KSKTLFSLMOYSEBRIHFTPEKQNNPNPSLIXPIPTOLLG-RTHATATGLRKVRRELFN 224
DB	206	PTKTQVGLTQYANNPRVFNLTNYTKTEEMIVATVSQTSQYGGDLTNTFGAIQVARKYAYS 265
QY	225	ITNGARKNAFKILFLLTIDGKFGPDGLGYEDVIEPLDRGVRVYIGV-----GDAPRSEK 279

Query Match 17.9%; Score 1054; DB 1; Length 1181;  
Best Local Similarity 26.7%; Pred. No. 2.4e-62;  
Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;  
QY 1 FNLDTENAMTPQ-ENARGFGQSVQL---QGRVVVVGAPQEIIVAAQNRGLSYQC--DYST 54  
D 30 YNVLPEAKITFSPSEQFGYVQVQINPKGNWLLVGSFWSGPPENMGVYKCPVDLST 89  
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSAAITSPQLLACGPTVHOTCSENYVUGLC 107  
D 90 ATCEKLNQTSIPNTWETKMTMSJGLILTRNMGTFGGFLTCGLMAQCGQNOYTTGVC 149  
QY 108 FLFGSNLRQOPQKPFALRGCPQEDSDIAFLIDGSGSIIPHDPFRKKEWVSTVMEQLK-- 165  
D 150 SDISDPF-QLSASFSPATQCPSL-IDVVVVCDESNSIYPWD--AVKNLEKXVQGLDIG 205  
QY 166 KSKTLFSLMOYSEFRIHTFKFQNNPNRSLIKXITQLLG-RYTATGLRKYVRELFN 224  
D 206 PTKTQVGLIQYANNPRVFNLTNTYKTEEMIVATSOYSQYGGDLNTWFGAIQYARKIAYS 265  
QY 225 ITNGAKNAFKILFLITDGEKFDPLGYEDVPELDREGVIRVIGV-----GDAPRSEK 279

RESULT 15  
ITAG HUMAN  
ID ITAG HUMAN STANDARD; PRT; 1167 AA.  
AC O75578; Q3UH28;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-10 precursor.  
GN ITGA10.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Articular chondrocytes;  
RX MEDLINE=98352078; PubMed=9685391;  
RA Camper L., Hellman U., Lundgren-Aakerlund E.;  
RT "Isolation, cloning, and sequence analysis of the integrin subunit  
RT alpha10, a beta1-associated collagen binding integrin expressed on  
RT chondrocytes";  
RL J. Biol. Chem. 273:20383-20389 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Endothelial cells, and Heart;  
RX MEDLINE=20169197; PubMed=10702680;  
RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,  
RT Wang S.-X., Langley R., Krissansen G.W.;  
RT "The integrin alpha10 subunit: expression pattern, partial gene  
RT structure, and chromosomal localization";  
RL Cytogenet. Cell Genet. 87:238-244 (1999).  
CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10  
CC ASSOCIATES WITH BETA-1.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Widely expressed with highest expression in  
CC muscle and heart. Found in articular cartilage.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF074015; AAC31952.1; -  
CC EMBL; AF112345; AAF21944.1; -  
CC EMBL; AF172723; AAF61638.1; -  
CC HSSP; P17301; 1A0X.  
CC Genew; HGNC:6135; ITGA10.  
CC MIM; 604042; -  
CC GO; GO:0008305; C:integrin complex; TAS.  
CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
CC GO; GO:0005518; F:collagen binding; TAS.  
CC GO; GO:0007160; P:cell-matrix adhesion; TAS.  
CC InterPro; IPR000413; Integrin\_alpha.  
CC InterPro; IPR002035; VWFA.  
CC Pfam; PF01839; FG-GAP; 3.  
CC Pfam; PF00092; vwa; 1.  
CC PRINTS; PR01185; VWFA DOMAIN.  
CC PRINTS; PR00453; VWFA DOMAIN.  
CC SMART; SM00191; Ict\_alpha; 4.  
CC SMART; SM00327; VWFA\_1.  
CC PROSITE; PS00242; INTEGRIN\_ALPHA; FALSE\_NEG.  
CC PROSITE; PS0234; VWFA; 1.  
CC Signal; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
CC Signal; Repeat; Calcium; Magnesium.  
FT SIGNAL 1 22  
FT CHAIN 23 1167 INTEGRIN\_ALPHA-10.  
FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1123 1145 POTENTIAL.  
FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 38 97 FG-GAP 1.

FT REPEAT 167 350 ? PG-GAP 2.  
FT DOMAIN 365 427 VWFA.  
FT REPEAT 428 482 FG-GAP 3.  
FT REPEAT 483 545 FG-GAP 4.  
FT REPEAT 546 605 FG-GAP 5.  
FT REPEAT 606 660 FG-GAP 6.  
FT REPEAT 661 724 FG-GAP 7.  
FT DOMAIN 725 1140 POLY-LEU.  
FT CA\_BIND 494 502 POTENTIAL.  
FT CA\_BIND 503 558 POTENTIAL.  
FT CA\_BIND 559 620 POTENTIAL.  
FT DISULFID 76 86 BY SIMILARITY.  
FT DISULFID 666 675 BY SIMILARITY.  
FT DISULFID 681 736 BY SIMILARITY.  
FT DISULFID 789 795 BY SIMILARITY.  
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 844 844 I -> L (IN REF. 2).  
FT CONFLICT 909 909 G -> V (IN REF. 2).  
FT CONFLICT 926 926 E -> D (IN REF. 2).  
SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AE80 CRC64;  
Query Match 17.9%; Score 1051.5; DB 1; Length 1167;  
Best Local Similarity 28.7%; Pred. No. 3,4e-62;  
Matches 351; Conservative 199; Mismatches 491; Indels 183; Gaps 45;  
QY 1 FNLDTENAMTEQENARG-FGOSVVLQGSRR---VVVGAPOSIIVANORGSLYQC----- 50  
DB 23 FNLDEHPRLPGGPPEAEFGYSLVHQVGGGQWMLVGAAPWDGPGDRGDVYRCPVGAH 82  
QY 51 -----DYSTG-SCEPIRLQVPVEAVNMISLGLSLAATSPPLACQPTVHQTCE 99  
DB 83 NAPCAKHLGIDYQLGNSSHP-----AVNMHLMGSLLETGDDGGFWACAPLMSRACGS 134  
QY 100 NTYVKGICFLFGSNLRQOPQKFPALRGCPQEDSDIAFLIDGSGSIIPHDFRKNEN-V 157  
DB 135 SVFSSGICARVDASFPQGS LAPTAQR-CPTY-MDVVIVLDGSSNIYP-----MSEV 184  
QY 158 STVMEQL-----KSKTFLSLMOYSEEFRIHFTFKBPQNNPNRSLIKPTQLLGR-TH 210  
DB 185 QTFLLRLVGLFIDPEQIQVGLVOYGSPVHWSIGDPFKBEVVVRAAKNLSRREGRTK 244  
QY 211 TATGLRKVVRELFNITNGARKNAFKILFLTDGKFP-GDPLGYEDVIFELDREGVIRVI 269  
DB 245 TAAQIMVACTEGFSQSHGGRPEAARLLVWVTDGSHDGEEL--PAALKACAGAVTRYGI 302  
QY 270 GV-GDAPRSEKS---ROELNTVASKPPDRHVFOINNFALKTIONOLREKIFAIEGTOT 324  
DB 303 AVLGHYLRQRDPSPLREITIASDPDERFFNVYDEAALTDIVDALGDRIFGLEGSHA 362  
QY 325 GSSSFHEMSQSGFSAAITNSGPELLSTVGSYDWDAGGVFLYTSKEKSTPIINTRVDS--- 381  
DB 363 ENESFGLMEQIGFSTHRLKXGILFGWVGAYDWGSGVLWLEGGHRLPPPRMALEDEFPP 422  
QY 382 --DMDAYLVGYA-AATILNRVQSVILGAPRYOHVIGLVAMER-ONTGMWESNANVKGTOI 437  
DB 423 ALQNHAAVLGYSVSSMLLRGRRFLPSGAPRFRHGRKVIAFQKKDGVAVRVAQSLQGGQI 482  
QY 438 GAYFCASLCSDVIDNSGSTDVLVLGAPHYY--EOTRGQGVSVCPPLPRQORARWQCDVAVI 495  
DB 483 GSYFGSELCPDTRDGTDTVLLVAAPMFLGPQNKETGRVYVYLV--QQOSSLTLQGTIQ 540  
QY 496 GEQGGPWGRFGAALTVLGDVNGDKLTVAIGA PGPEENRNGAVYLFHGTSGGISPSHSOR 555

Db 541 PEPQD-ARPGFANGALPDLNODGFADVAVGAPLEDHQGALYLHGTQ-SGVREHPAQR 598  
Qy 556 IAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQHVLILLRSQPLRVKAIMENPRE 615  
Db 599 IAAASMPHALSYFGRSVDRGLDLDGDDLDVAVGAQAAILLSSRPVHLTPSLEVTPOA 658  
Qy 616 VARNVFECNDQVVKGEAG-EVRVCLHVQKSTRDRLEGQIOSVVVYDIALDSGRPHSR 673  
Db 659 ISVVQDCCR--RGQAVCLTALCFQVTSRTPGRWDH---QFYMRFTASLDEWTAGAR 712  
Qy 674 AVNET--KNSTRQTOVLGLTCTETLKLQLPNCIEDVSPVIVLRNLFNLVGTPLSARG 731  
Db 713 AAPDGSQRLSPRRLRSVG-NVTCEQLHFNULD-TSDYLRPVALVTWFDLNTTKPG-- 768  
Qy 732 NLRPVLAEQAQRLFTALPPERNCGNDNICODLSITFSFMSLDC-----LVVGGPR 783  
Db 769 ---FVLNBSPTSIOKLVPFSKDCGPDNECVTLVLQ---VMDIRGSRKAPFVVRGGR 822  
Qy 784 EFNVTVVRNGEDSVYRQVTFPPLDLVYKVKSTLQNRQSRWRLACESASSTEVSGA 843  
Db 823 KVLVSTTLNKENAYNTSLSIIFSRNL---HLASLTPQR-ESPIKVECAAPSA----- 872  
Qy 844 LKSTSCSINHPIFPENSEVTENITEDVDSKASLG---NKLL-----LKANVTSENNMPT 895  
Db 873 -HARLCSVGHVPFQTKAKVTFLLFEFSCSLUSQVFGKLTASSDSLERNGLQENTACT 931  
Qy 896 NKTEFQLELFPVAVYVMTVSHGVSTKYNFTASENTSRYMCHQY----- 940  
Db 932 -----SAYIQYEPH-----LLFSSESTLHRYEVHPYGTLPVGPGRPKTTL 972  
Qy 941 QVSNLG---QRSLPISLVFLVP-----VRLNQTVIWRDPQVTFSENLSSTCHTKER 988  
Db 973 RVQNLGCVYVSGHIIISA--LLPAVARGNYPUSLSQVI-----TNNASCIQVNLTE 1021  
Qy 989 LPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIPFGIQBEFNATLKGNLSDWDYIKTSHNH 1048  
Db 1022 PPGPPVHPELOHTNRLNGSNTQCVVRCHLGOLANGTEVSVCGLRLVNEHERRAKFS 1081  
Qy 1049 LLIVSTAELLFNDVSFTLLPGQGAFTVRSQTEKVBEPPEVNPPLIVGSSVQGLLLALI 1108  
Db 1082 LTVVSTFELGTERGSLQLTEASRWSESLEEV-VQTRPILISLWILIGSVLQGLLLALL 1140  
Qy 1109 TAALYKLGPF-----KROYK 1123  
Db 1141 VFCLWKLGPFAHKKIPEERKREK 1164

Search completed: June 7, 2004, 17:12:47  
Job time : 14.9719 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 16:56:30 ; Search time 39.4889 Seconds

(without alignments)  
9084.693 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFQ.....PKRQYKDMSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.rodent.\*
- 12: sp.virus.\*
- 13: sp.vertebrate.\*
- 14: sp.unclassified.\*
- 15: sp.virus.\*
- 16: sp.bacteriap.\*
- 17: sp.archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4371	74.3	1151	11 Q9J130	Q9j130 rattus norv
2	3891.5	66.1	1036	11 Q8CA73	Q8ca73 mus musculus
3	3759.5	64.6	1920	6 Q28984	Q28984 sus scrofa
4	3470	59.0	1169	4 Q81VAG	Q81vag homo sapien
5	3310.5	56.3	1169	11 Q9QXH4	Q9qxh4 mus musculus
6	3236.5	55.0	1161	11 Q9QYB7	Q9gye7 rattus norv
7	1523.5	25.9	1161	11 Q9WTV4	Q9wtv4 mus musculus
8	1513	25.7	1160	11 Q9R200	Q9r200 mus musculus
9	1395	23.7	1196	13 Q98TF1	Q98tfl cyprinus ca
10	1350.5	23.0	1086	4 Q96HB1	Q96hbl homo sapien
11	1344.5	22.9	1187	13 Q98TF0	Q98tfo cyprinus ca
12	1269	21.6	927	6 Q8HZV0	Q8hzv0 bos taurus
13	1154.5	19.6	1167	11 Q88340	Q88340 rattus norv
14	1109	18.8	1167	11 Q88341	Q88341 rattus norv
15	1043.5	17.7	1188	11 Q7QC3	Q7tc3 mus musculus
16	1041	17.7	1171	13 Q42094	Q42094 gallus gall

17	1032	17.5	1038	11 Q8BS01	Q8bs01 mus musculus
18	1005	17.1	895	11 Q9WUF8	Q9wuf8 mus sp. itg
19	998.5	17.0	1160	6 Q8MK74	Q8mkf4 felis silve
20	865	14.7	348	4 Q8TES5	Q8tes5 homo sapien
21	860	14.6	1332	5 Q3BPQ8	Q3bpq8 halocynthia
22	796	13.5	205	11 Q63001	Q63001 rattus norv
23	753.5	12.8	780	13 Q06271	Q06271 xenopus lae
24	738	12.5	823	4 Q8WY18	Q8wy18 homo sapien
25	686.5	11.7	823	11 Q8CE84	Q8ce84 mus musculus
26	669	11.4	1032	11 Q61989	Q61989 mus musculus
27	644.5	11.0	1033	6 Q9BGU3	Q9bgu3 bos taurus
28	643	10.9	1036	11 Q91YD5	Q91yd5 mus musculus
29	623.5	10.7	1474	5 Q86G87	Q86g87 pseudoplusi
30	628.5	10.7	257	11 Q8C270	Q8c270 mus musculus
31	619.5	10.5	1041	5 Q9UB90	Q9ub90 lytechinus
32	614.5	10.4	1041	5 Q76378	Q76378 lytechinus
33	580	9.9	1034	13 Q98T77	Q98tt7 gallus gall
34	579.5	9.8	1054	5 Q9UGS1	Q9ugs1 strongyloce
35	555.5	9.4	1053	11 Q80YD5	Q80yd5 mus musculus
36	550	9.3	1033	13 Q42598	Q42598 xenopus lae
37	546	9.3	1036	6 Q7YRP8	Q7yrp8 equus cabal
38	534	9.1	1016	13 Q91779	Q91779 xenopus lae
39	530	9.0	974	11 Q924W2	Q924w2 rattus norv
40	529	9.0	1073	11 Q8CC06	Q8cc06 mus musculus
41	527.5	9.0	1119	5 Q86G88	Q86g88 pseudoplusi
42	526	8.9	1047	6 Q9WZD6	Q9wzd6 bos taurus
43	525.5	8.9	1007	6 Q9GK48	Q9gk48 bos taurus
44	522.5	8.9	1132	11 Q80Z18	Q80z18 mus musculus
45	512.5	8.7	1034	6 Q9TUN4	Q9tun4 oryctolagus

#### ALIGNMENTS

#### RESULT 1

Q9J130 PRELIMINARY; PRT; 1151 AA.  
 AC Q9J130; TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE Integrin beta 2 alpha subunit.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fathallah D.M. Sr., Zerria K. Jr.;  
 RT "Cloning of the rat CD11b cDNA sequence."  
 RL Submitted (MAY-2000) TO THE EMBL/GenBank/DBJ databases.  
 DR EMBL; AF268593; AAF81280.1; -  
 DR HSP; P11215; 1BHQ.  
 DR GO; GO:0008305; C:integrin complex; IEA.  
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PRO1185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Intc\_alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.3%; Score 4371; DB 11; Length 1151;  
 Best Local Similarity 72.6%; Pred. No. 1.9e-313;  
 Matches 826; Conservative 150; Mismatches 159; Indels 2; Gaps 2;

```
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQBIIVAAHQBSGLYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFOENARGFGQSVVQLQGSRRVVGAPQBIIVAAHQBSGLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNNVSLGLSLAAATTPQQLACGPTVHQTCSENTYVKGCLFGLGSLNRQOPQK 120
Db 77 PLQVPEAVNNVSLGLSLAAATTPQQLACGPTVHQTCSENTYVKGCLFGLGSLNRQOPQK 136
QY 121 FPEALRCQDSDIAFLIDGSGSIIPHDPRFMKEWSTVMBOLKSKTFLSLMOYSEBF 180
Db 137 FPEALRCQDSDIAFLIDGSGSIIPHDPRFMKEWSTVMBOLKSKTFLSLMOYSEBF 196
QY 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRXVRELEWITNGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRXVRELEWITNGARKNAFKILFL 256
QY 241 TDGSKFGDPLGYEDVPELDRGVRVIVGVGDAFRSEKSRQELNVTASKPPDRHVQIN 300
Db 257 TDGSKFGDPLGYEDVPELDRGVRVIVGVGDAFRSEKSRQELNVTASKPPDRHVQIN 316
QY 301 NFALKTIONLREKIPIAETGQTGSSSPHEHMSQEGFSAATNSGFLSTVGSYDNAG 360
Db 317 NFALKTIONLREKIPIAETGQTGSSSPHEHMSQEGFSAATNSGFLSTVGSYDNAG 376
QY 361 GVFLYTSKXSTFTNTRVDSMDNDAYLGYAAIIENRVOSVLGAPRYOHIGLVAMFR 420
Db 377 GAFLYPSKDKASFTNTRVDSMDNDAYLGYAAIIENRVOSVLGAPRYOHIGLVAMFR 436
QY 421 QNTGWESNANVKTQIGYFAGSLCSVDVDSNGSTDLVLIGAPHYYBOTRGQGVSCPL 480
Db 437 QNTGWESNANVKTQIGYFAGSLCSVDVDSNGSTDLVLIGAPHYYBOTRGQGVSCPL 496
QY 481 PRGORARQWQDAVLYGQGGQFMRFGAALTVLGVDVNGDKLTDVAIGAPGEDNKRGAIVLF 540
Db 497 PRGORARQWQDAVLYGQGGQFMRFGAALTVLGVDVNGDKLTDVAIGAPGEDNKRGAIVLF 555
QY 541 HGTSGSGISPSHQRINGSLRLOVFGQSLGGQDLTDGGLVDTLTVGQGHVLLRQ 600
Db 556 HGTSGSGISPSHQRINGSLRLOVFGQSLGGQDLTDGGLVDTLTVGQGHVLLRQ 615
QY 601 PVLRVKAIMEFNEPREVARNVFCNDQVKGKAGEVRVCLVHVKSTRDLRREGQIOSVT 660
Db 616 PVLRVKAIMEFNEPREVARNVFCNDQVKGKAGEVRVCLVHVKSTRDLRREGQIOSVT 675
QY 661 YDLALDSGRPHSRAVNETKSTRQTVGLQTCETLKLQPCNCEIDPVSIVRLNPF 720
Db 676 YDLALDSGRPHSRAVNETKSTRQTVGLQTCETLKLQPCNCEIDPVSIVRLNPF 735
QY 721 SLVGTPLSAPGNLRPVLAEADQRLFTALPPEKNCQNDNICQDLSITPFSNLDCLVVG 780
Db 736 SLVGTPLSAPGNLRPVLAEADQRLFTALPPEKNCQNDNICQDLSITPFSNLDCLVVG 795
QY 781 GPREFNVTVVRNDGDSYPTQVTFPDLDSLRKYSTLQNSQSRWLAACESASTEV 840
Db 796 GPREFNVTVVRNDGDSYPTQVTFPDLDSLRKYSTLQNSQSRWLAACESASTEV 854
QY 841 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTNKTFF 900
Db 855 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTNKTFF 914
QY 901 QLELPVKYAVMYVTSHGVSTKYLNTFASNTSRVMOHQYQVSNLQORSLPLSLVLPV 960
Db 915 QLELPVKYAVMYVTSHGVSTKYLNTFASNTSRVMOHQYQVSNLQORSLPLSLVLPV 974
QY 961 RLNQTVDRPQVTFENISSTCHTERLPSSHDLAEARKAPVNCSTAVCQRQCDDLP 1020
Db 975 RLNQTVDRPQVTFENISSTCHTERLPSSHDLAEARKAPVNCSTAVCQRQCDDLP 1034
QY 1021 FPGIOEFNATLKNLSFDWYIKTSENHLLIVSTABILFNDSYFTLLPQCGAFVRSQTE 1080
Db 1035 FPGIOEFNATLKNLSFDWYIKTSENHLLIVSTABILFNDSYFTLLPQCGAFVRSQTE 1094
QY 1081 KVEPPEVPEPLPLIVGVSSVGLLILLALITAAALKLGFRRQYKDMMSGEGPFGAEPQ 1137
```

```
Db 1095 KVEPPTVHPVPLIVGVSSVGLLILLALITAGLYKLGFFRRQYKDMNTEAGGQDGPQ 1151
RESULT 2
Q8CA73 PRELIMINARY; PRT; 1036 AA.
AC Q8CA73;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Integrin alpha M.
GN ITGAM OR P730045J24RIK.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=22354683; PubMed=12466851;
RA THE PANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT -Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.;
RL Nature 420:563-573 (2002).
DR EMBL; AK039444; BAC30350.1; -.
DR PIR; PT0572; PT0572.
DR PIR; PT0633; PT0633.
DR PIR; PT0697; PT0697.
DR MGD; MGI:96607; Itgam.
DR GO; GO:0008305; C: Integrin complex; IEA.
DR GO; GO:0004895; P: cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P: cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
SQ SEQUENCE 1036 AA; 115126 MW; 4F4F39BF1E188B77 CRC64;
Query Match 66.1%; Score 3891.5; DB 11; Length 1036;
Best Local Similarity 66.0%; Pred. No. 4,3e-278;
Matches 751; Conservative 125; Mismatches 143; Indels 119; Gaps 2;
```

```
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQBIIVAAHQBSGLYQCDYSTGSCPEI 60
```

```
Db 17 FNLDTEHPMTFOENARGFGQSVVQLQGSRRVVGAPQBIIVAAHQBSGLYQCDYSTGSCPEI 76
```

```
QY 61 RLQVPEAVNNVSLGLSLAAATTPQQLACGPTVHQTCSENTYVKGCLFGLGSLNRQOPQK 120
```

```
Db 77 PLQVPEAVNNVSLGLSLAAATTPQQLACGPTVHQTCSENTYVKGCLFGLGSLNRQOPQK 136
```

```
QY 121 FPEALRCQDSDIAFLIDGSGSIIPHDPRFMKEWSTVMBOLKSKTFLSLMOYSEBF 180
```

```
Db 137 FPEALRCQDSDIAFLIDGSGSIIPHDPRFMKEWSTVMBOLKSKTFLSLMOYSEBF 196
```

```
QY 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRXVRELEWITNGARKNAFKILFL 240
```

```
Db 197 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRXVRELEWITNGARKNAFKILFL 256
```

```
QY 241 TDGSKFGDPLGYEDVPELDRGVRVIVGVGDAFRSEKSRQELNVTASKPPDRHVQIN 300
```

```
Db 257 TDGSKFGDPLGYEDVPELDRGVRVIVGVGDAFRSEKSRQELNVTASKPPDRHVQIN 316
```

```
QY 301 NFALKTIONLREKIPIAETGQTGSSSPHEHMSQEGFSAATNSGFLSTVGSYDNAG 360
```

```
Db 317 NFALKTIONLREKIPIAETGQTGSSSPHEHMSQEGFSAATNSGFLSTVGSYDNAG 376
```

```
QY 361 GVFLYTSKEKSTINMTRVDSMDNDAYLGVAAILLRNRVQSVLVGAPRYQHIGLVAMFR 420
DB 377 GAFLYTSKOKVITINTRVDSMDNDAYLGVASAVILRRNRVQSVLVGAPRYQHIGLVAMFR 436
QY 421 QNTCMFESNANWGTQIGAFGASCLSDVDNMGSDLVGAPRYQHIGLVAMFR 480
DB 437 ENFGTWEPHTSIKG-----450
QY 481 PRQORARWQDAVLYGEGQGWGRFGAALTVLGDNVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 451 -----450
QY 541 HGTSGSGISPHSRIAGSKLSPRLOYFGOSLSGGDLTMDGLVDTVGAQGHVALLRSQ 600
DB 451 -----SORLIGAHFSPLOYFGOSLSGGDLTMDGLVDTVGAQGHVALLRSQ 498
QY 601 PVLVRKALMEFNRPREVARNVFCNDQVVGKEAGEVRCVHLVQKSTRDLREGIQSVYT 660
DB 499 PVLRLBATEFSPKPKVARSVFACQEQVLKXKADAGEVRCVLRVKNKTKRLREGIQSVYT 558
QY 661 YDLALDSGRPHSAVNETKSTRROTGLTCTCETLKLQLPNCIEPVPSPVILRLNP 720
DB 559 YDLALDSGRPHSAVNETKSTRROTGLTCTCETLKLQLPNCIEPVPSPVILRLNP 618
QY 721 SLVGTPLSFGNLRPVLAEDAORLFTALPFFERNKCGNDNICQDLSITFSFMSLDCLVVG 780
DB 619 TLVGEPLRSFGNLRPVLAEDAORLFTALPFFERNKCGNDNICQDLSITFSFMSLDCLVVG 678
QY 781 GRPEENVTVVRNDGDSYRTQTFPPFLDLSVRKYSTLQXQBSQSWRL-ACESASSTE 839
DB 679 GPQDFNMSVTLRNDGDSYRTQTFPPFLDLSVRKYSTLQXQBSQSWRL-ACESASSTE 738
QY 840 VSCALSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANKVTSNNPRTNKTE 899
DB 739 GHCAKASTTWINHPIFPANSEVFNITFDVDSKASLGKLLKANKVTSNNPRTNKTE 798
QY 900 FQLELPVKYAVVWTSYHGVSTKYNFTASENTSRVQVQVQVQVQVQVQVQVQVQV 959
DB 799 FQLELPVKYAVVWTSYHGVSTKYNFTASENTSRVQVQVQVQVQVQVQVQVQVQV 858
QY 960 VRLNQTWIDRPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCRIOQCDI 1019
DB 859 VQINNVTVWHDPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCRIOQCDI 918
QY 1020 PFGIOBEPNATLKNLSFDWYIKTSHNHLIIVSTABILLFNDSVFTLLPQCGAFVRSQTE 1079
DB 919 PSNTQIEIFNVTLKGNLSFDWYIKTSHNHLIIVSTABILLFNDSVFTLLPQCGAFVRSQTE 978
QY 1080 TKVEPEVNPPLIVGSSVGLLALITAMLYKLGFKEKROYEDMWSGEGPPGAPQ 1137
DB 979 TKVEPEVNPPLIVGSSVGLLALITAMLYKLGFKEKROYEDMWSGEGPPGAPQ 1036

RESULT 3
Q28984
ID Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN CD11B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR ENBL; U00072; AAB16869.1; -.
DR HSSP; P11215; 1BHQ.
```

```
DR GO:0008305; C:integrin complex; IEA.
DR GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON_TER 1
FT SEQUENCE 920 AA; 102440 MW; E96CC51B350DD5AC CRC64;
SQ
Query Match 64.6%; Score 3799.5; DB 6; Length 920;
Best Local Similarity 78.8%; Pred. No. 2.2e-271;
Matches 726; Conservative 84; Mismatches 110; Indels 1; Gaps 1;
QY 118 POKFPEARLGCQEDSDIAFLIDGSGSIIIPDPRMKEMVSTVMSQLKSKTLPSLMQYS 177
DB 1 POKFPEARLGCQEDSDIAFLIDGSGSIIIPDPRMKEMVSTVMSQLKSKTLPSLMQYS 60
QY 178 EEFRIHPTFKFQNNPNSRIKPIITQLLGRTHATGLRKVVRLEFNITNGARKNAFKIL 237
DB 61 EDFYTHFTFNDPKRNPSEKLLVPEIRQLLGRTHATGIRKVVYRELPHSKSGARENAKIL 120
QY 238 FLITDGEKGPGLGVEDVITELDRGEGVRYVGVGDADFSEKSRQELNATVASKPRDHVF 297
DB 121 VVITDGEKGPGLGVEDVITELDRGEGVRYVGVGDADFSEKSRQELNATVASKPRDHVF 180
QY 298 QINNPEALKTKQNLREKIFAIETGQTSSSSFFHEMSQEGFSAAITSNGLPSTVGSYD 357
DB 181 QVNNPEAVKTIQNLQKTEFALGEGTQTSSTSFCEMSEGEFSAITSNGLPSTVGSYD 240
QY 358 WAGGVFLTSYKSTKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSVLVGAPRYQHIGLV 417
DB 241 WAGGVFLTSYKSTKSTFNNTRVDSMDNDAYLGAAAAIILNRVQSVLVGAPRYQHIGLV 300
QY 418 MFRONTGMWSENANVKGTOIGAYFGASLCSDVDNMGSDTLVIGAPHYEOTRGQVSV 477
DB 301 MFKQNSGAWENKADIKSGISYFGASLCSDVDNMGSDTLVIGAPHYEOTRGQVSV 360
QY 478 CPLRQGRARWQDAVLYGEGQGWGRFGAALTVLGDNVNGDKLTDVAIGAPGEEDNRGAV 537
DB 361 CPLRQGRARWQDAVLYGEGQGWGRFGAALTVLGDNVNGDKLTDVAIGAPGEEDNRGAV 419
QY 538 YLFHGTSGSGISPHSRIAGSKLSPRLOYFGOSLSGGDLTMDGLVDTVGAQGHVALL 597
DB 420 YLFHGTSGSGISPHSRIAGSKLSPRLOYFGOSLSGGDLTMDGLVDTVGAQGHVALL 479
QY 598 RSQPVLRVYKATMEFNPREVARNVFCNDQVVGKEAGEVRCVHLVQKSTRDLREGIQS 657
DB 480 RSQPVLRVYKATMEFNPREVARNVFCNDQVVGKEAGEVRCVHLVQKSTRDLREGIQS 539
QY 658 VVTVDLALDSGRPHSAVNETKSTRROTGLTCTCETLKLQLPNCIEPVPSPVILRL 717
DB 540 VVTVDLALDSGRPHSAVNETKSTRROTGLTCTCETLKLQLPNCIEPVPSPVILRL 599
QY 718 LNFSLVGTPLSFGNLRPVLAEDAORLFTALPFFERNKCGNDNICQDLSITFSFMSLDC 777
DB 600 LNFSLVGTPLSFGNLRPVLAEDAORLFTALPFFERNKCGNDNICQDLSITFSFMSLDC 659
QY 778 VVGPPREFNVTVVRNDGDSYRTQTFPPFLDLSVRKYSTLQXQBSQSWRL-ACESAS 837
DB 660 VVGPPREFNVTVVRNDGDSYRTQTFPPFLDLSVRKYSTLQXQBSQSWRL-ACESAS 719
QY 838 TEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANKVTSNNPRTNK 897
DB 720 TEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANKVTSNNPRTNK 779
QY 898 TEFQLELPVKYAVVWTSYHGVSTKYNFTASENTSRVQVQVQVQVQVQVQVQVQV 957
```



```

DR HSSP; P11215; 1BHQ.
DR MGD; MGI:96609; Itgax.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
DR Integrin.
KW INTEGRIN.
SQ
  Query Match      56.3%; Score 3310.5; DB 11; Length 1169;
  Best Local Similarity 56.6%; Pred. No. 4.4e-235;
  Matches 645; Conservative 173; Mismatches 303; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAHQSGLSYQCYSTGSCPEI 60
Db 20 FNLDAEKLTHFMDGAEFGHSLQVLDSSWVVGAPKEIKATNQIGGLYKGYHTGNCPEI 79
QY 61 RLQVPEAVNMSIGLSLAATTSPOLLACGPTVHOCSENTYVKGCLFLGFSNLRQOQK 120
Db 80 SLQVPEAVNISLGSLSAATNPMLLACGPTVHHTCRENIYLTGLCLFLSSFFQS-QN 138
QY 121 FPEALRGCPQEDSDIAFLDGGSGSIIPHDPRKEMWSTVMBQLKKSKTLFSLMOYSEEF 180
Db 139 FPTAQCECPKQDQIVFLDGGSGSIISDTDFXWLDVFXAVMSQLQRPSTFSLMQFSYF 198
QY 181 RHFTFKFQNNPNRSLIKPI:TOLLGRHTATGLRKVVRELFNITNGARKNAFKILFL 240
Db 199 RVHFTFNNPISSTPLSLGVRQLRGYTTTASAKHVTITLFTQSGARQDQATKVLVI 258
QY 241 TDGEKFGDPLGYEDVIPELDRGSEYRVYGVGDGAFRSKRSQRLNTVASKPPDRVFGQIN 300
Db 259 TDGRKQGNLSYDSVTPMAEASIIRYAIGVCKAFVNEHSKQELAKIASMSEHVFSVE 318
QY 301 NFEALKTONQUREKIFAIEGHTQSSSFHEMSQEGFSAAITNGPLSTLVGSDYDAG 360
Db 319 NFDALKDIEHQLEKIPAEIGETFTSSSTFELMSQEGFSAVFTPDGVLGAVGFSWSG 378
QY 361 GVFLYTSKSKSTFINNTRVDSMDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
Db 379 GAFLYPSNMRPTFINNSQENEDRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIFT 438
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYEQTRGGQVSVCP 480
Db 439 QSSREHWRPKSEVRGTQIGSYFGASLCSVDMDRSGSTDVLVIGPHYHYEHTRGQVSVCPM 498
QY 481 PRGQARWOCDAVLGECQPGRGGAALTIVLGDVNGDKLTDVAIGAGEEDNRCAYVLF 540
Db 499 P-GVGRHHCQYTLHGEQHPGRGAALTIVLGDVNGDSLADVALGAPGEENRGAVYIF 557
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMQGLVDLTVGAQHVLILRSQ 600
Db 558 HCAQRQDIAPSPQRIASQIPSRIOYFGQSLSGGQDLTRDGLVDLAVGSKGRVLLNTR 617
QY 601 PVLRYKAMEPNPREVARNVFCNDVYKGEAGEVRVCLHVQKSTRDLREGQIQSVTT 660
Db 618 PLRVSPTVHFPASIRSVFECQVAPQEQVLTSDATVCLHSHSPKQL--GDLRSVTI 675
QY 661 YDLALDSGRPHSRVAFNTEKSTRQTQVLGTLQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 676 FDLALDHGRSLSTRAIFKTKTRALTRVTLGLNKECESVKLLLPACVEDSDVPTILRNP 735
QY 721 SLVGTPLSAFNGLRPVLEADAQRLTALFPPEFKNGCNDNICQDDLSTFSPMSLDCLVVG 780

736 SLVGPVSISSLQNLQMLAVDDQTYFTASLPFEKNGADHICQDDLSVVFPGFDLTKLVWG 795
781 GPRFNVTVTVNRNDCGDSYRTQVTFFPPLDLISYKXVSTLQ-----NQRQGRSWR 829
796 SDLELVNDVTVSNDGDSYGTIVTLFYPVGLSPRRVABGQVFLRKCKEQDQWRQRGHSILH 855
830 LACESASSTEVSGALKKSTSCSINHIPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSE 889
856 LMCD--STPDRSQGLWSTSCSRHVIPIFGSQMTFLVTFDVPKABLDRLLLRARVGSE 913
890 NMPRTNKTETQLLPVKYAVYVMVTSRGSVTKYKLNFTASB-NTSRVNMHQHVQVSNLGR 948
914 NNVPCTPKTTFQLLPVKYAVYTWISSDDQTKYLNFTSEKETSVEHRFQVNNLQGR 973
949 SLPISLVFLVPLRNQTVIMDRPQVTFSENLSSTCHTKERLPSSHDFIAELRKAPVNVCS 1008
974 DVPVSINFVPIELKGEAVW-TVMVSHPNQLTCYRNRLKPTQFDLLTTHMOKSVLDCS 1032
1009 IAVCORIOCDIPFGIOEFNATLKGNLSPDWYIKTSHNLLIIVSTAEILFNDSVFTLLP 1068
1033 IADCLHLECDIPSGILDLYFLKGNLSFGWISQTLQKKVLLISEAIFNTSVYSQLP 1092
1069 GQGAFFVASQTKVPEPEVPNPPLIIVGSSVGGLLALLALITAALYKLGFFKRYKQKMMSE 1128
1093 QGEAFLEAQTATVLEMYKVNHPVPLIVGSSVGGLLALLAITAALLYKAGFFKRYKEMLEE 1152

RESULT 6
QYQYEV7 PRELIMINARY; PRT; 1161 AA.
AC QYQYEV7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
OS Alpha D integrin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., Vandervieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin."
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
DR Integrin.
KW INTEGRIN.
SQ
  Query Match      55.0%; Score 3236.5; DB 11; Length 1161;
  Best Local Similarity 57.5%; Pred. No. 1.3e-229;
  Matches 648; Conservative 163; Mismatches 303; Indels 13; Gaps 9;

QY 2 NLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQEIIVAAHQSGLSYQCYSTGSCPEI 61

```